ELSEVIER

Physica D 98 (1996) 343–378

PHYSICA D

Self-consistent Hamiltonian dynamics of wave mean-flow interaction for a rotating stratified incompressible fluid

Ivan Gjaja, Darryl D. Holm *

Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

Abstract

We derive a hierarchy of approximate models of wave mean-flow interaction (WMFI) by using asymptotic expansions and phase averages either in Euler's equations for a stratified rotating inviscid incompressible fluid, or in Hamilton's principle for these equations. Two small dimensionless parameters are used in these expansions. The first parameter is the ratio of time scales between internal waves at most wave numbers and the mesoscale mean flow of the fluid. This "adiabatic ratio" is small and is comparable to the corresponding ratio of space scales for the class of initial conditions that support internal waves. The other small parameter used in these expansions is the ratio of the amplitude of the internal wave to its wavelength. An application of Noether's theorem to the phase-averaged Hamilton's principle shows that the resulting equations conserve the wave action, convect a potentional vorticity and can, depending on the order of approximation, convect wave angular momentum. Passage from the phase- averaged Hamilton's principle to the Hamiltonian formulation brings the WMFI theory into the Lie–Poisson framework in which formal and nonlinear stability analysis may be applied. This framework also suggests a two-fluid model of the interaction of waves and mean flow similar to that for the superfluid and normal fluid components of liquid He^4 without vortices. We also discuss the relations of these results to the Charney–Drazin nonacceleration theorem, Whitham averaging, WKB stability theory, and Lagrangian-mean fluid equations for prescribed wave displacements.

Keywords: Internal waves; Phase averaging; Hamilton's principle; WKB approximation

1. Introduction

Waves in a hydrostatically stable stratified fluid whose restoring force is due to gravity are called "internal waves" These waves are ubiquitous in the ocean and have typical vertical displacements of about 10 m, with periods ranging from tens of minutes up to the rotational period. The associated horizontal fluid velocity is typically 0.05 m/s, with horizontal displacements of about 1 km. There are many sources of internal waves in the ocean, such as oscillations of the thermocline, tidally driven flow over topography and instabilities of energetic columnar eddies whose sizes tend to be nearly equal to the local internal Rossby deformation radius. For reviews of internal waves in the ocean and discussions of recent progress in the analysis of their dynamics, see, e.g. [20,21,38].

* Corresponding author. E-mail: dholm@lanl.gov

0167-2789/96/$15.00 Published by Elsevier Science B.V.
PII S0167-2789(96)00104-2

Since the time and space scales of the ocean's background mean motion which modulates the internal wave are large compared to its wavelength and period, the internal wave can be regarded as a wave train whose amplitude, frequency and wave vector vary slowly compared to its phase oscillations. Thus, there is a small "adiabatic"parameter $\epsilon \ll 1$, which is the ratio of time and space scales between the wave oscillations and the mean flow. This adiabatic parameter is given by $1/\epsilon = \omega T$, where ω is the wave frequency and T is the characteristic time scale of the mean flow (eddy turning time, for example), and *also* by $\epsilon = \lambda/L$, where λ is the wavelength of the internal wave and L is the characteristic scale of the mean flow. Under these circumstances, the trajectory of a Lagrangian fluid element X in a stably stratified flow decomposes naturally into the sum of a slowly varying mean displacement $\boldsymbol{x}$ and rapid fluctuation $\boldsymbol{\xi}$. Thus, one may write $\boldsymbol{X} = \boldsymbol{x} + \alpha\boldsymbol{\xi}(\boldsymbol{x}, t)$, where $\alpha \ll 1$ is the displacement of the fluid due to the wave divided by the wavelength, and $\boldsymbol{\xi}$ has zero mean. This wave mean-flow (WMF) decomposition suggests an approach based on asymptotic expansion in relative displacement amplitude α, combined with either a WKB analysis, or the method developed by Whitham [46,47], of averaging over the rapid – order $O(1/\epsilon)$ – phase of the wave in the fluid action principle in order to derive the equations for wave mean-flow interaction (WMFI) dynamics. This combined approach using asymptotic expansions and WKB analysis is standard for the WMFI problem, see, e.g., the review by Grimshaw [26] and related remarks in the review by Salmon [43].

In this paper, we derive an asymptotic hierarchy of new equations for WMFI dynamics in two different ways. In our first derivation, we substitute the assumed WMF decomposition into Euler's equations for an inviscid stratified rotating incompressible fluid and demand that the coefficients of mean and fluctuating terms vanish separately, to an appropriate order in ϵ and α. This approach is called the "direct expansion". Then, at each order of accuracy we rederive the *same* equations as those obtained from the direct expansion by using Hamilton's principle (HP) asymptotics, which in this case consists of substituting the WMF decomposition into HP for the unapproximated Euler dynamics, truncating the combined expansion in ϵ and α, and phase averaging before taking variations. The derivation by HP asymptotics uses an extension of Whitham's method of averaging in the action principle to produce a self-consistent description of ideal WMFI dynamics. The extension is to keep the Lagrangian labels (the initial coordinates) fixed in the phase averaging, regardless of whether their dependence is slow or fast. This approach parallels Whitham averaging for a finite dimensional system and has the advantage of allowing sufficiently general spatial dependence of the pressure and density fields in the WMFI equations to support internal waves.

While the derivation of the WMFI equations from Euler's equations by direct asymptotic expansion confirms the validity of their rederivation from HP asymptotics, the equations for WMFI dynamics resulting from expansion and phase averaging in HP generally contain at each truncation order certain additional higher-order terms which would be neglected at that order in using the direct asymptotic expansion approach. These "remainder" terms in the equations arising in HP asymptotics serve to provide exact conservation laws for WMFI arising from symmetries of HP. In fact, some of these symmetries are *created* in the process of asymptotic expansion and phase averaging in HP. Of course, the corresponding conservation laws are also approximately valid to a certain order in the direct expansion approach.

Retaining the higher-order terms in the equations arising from HP asymptotics at a given step provides exact conservation laws, but does not actually improve the order of accuracy of the equations. This is because not *all* the higher-order terms appear in the remainder arising from HP asymptotics, only those terms that provide the conservation laws due to symmetries of HP at a given order of its truncation. Furthermore, such symmetries may be produced in the process of making the approximations. Consequently, there is an "alternating feedback" available between the results obtained from direct asymptotic expansions in the equations of motion and results obtained from HP asymptotics when the unapproximated theory arises from HP. At each order in the expansion, HP asymptotics produces equations with a remainder containing the higher-order terms which provide exact conservation laws for the equations at that order. In the next iteration of the direct expansion, the rest of the higher-order terms in the equations are obtained, including the terms appearing in the HP remainder from the previous order. The next iteration of the direct expansion may then be used as potential input for another iteration of the HP asymptotics. Thus, the derivation by HP asymptotics does not produce anything which could not be obtained by the direct expansion method by going to a higher order. However, the derivation by HP asymptotics has the advantage of preserving the conservation laws due to symmetries in the unapproximated HP at each order, as well as the possibility of adding new conservation laws due to symmetries induced by averaging. Of course, averaging may be done over a variety of other parameters besides phase.

Having derived the WMFI equations both by direct asymptotic expansion in Euler's equations and by phase averaging in HP at a certain order, we pass to their Hamiltonian formulation via the Legendre transformation. In this Hamiltonian formulation, the energetics of WMFI is clear and its dynamics is portrayed as mean hydrodynamic motion in the Eulerian description evolving the on the slow space and time scales, at which the "rectified" effects of rapidly oscillating waves appear, roughly speaking, as though the waves were another fluid degree of freedom coupled to the mean flow. This degree of freedom is associated with the wave vector (defined as the slow spatial gradient of the phase) and the density of the conserved wave action. The wave action density is conserved, since it is the momentum density canonically conjugate to the phase of the wave field, which becomes an ignorable coordinate in the process of averaging. The product of the wave action density and the wave vector is the wave "pseudomomentum" density, which generates spatial shifts in wave quantities under the Poisson bracket operation while leaving mean-flow quantities invariant. Thus, the Hamiltonian formulation presented here provides a self-consistent picture of WMFI dynamics as a nonlinear fluid theory coupled to a subsystem degree of freedom whose coordinate field variable (the phase of the wave train) is made ignorable by averaging; so its canonically conjugate momentum density (the wave action density) satisfies an associated conservation law. This Hamiltonian formulation is a self-consistent alternative to previous investigations of WMFI.

WMF interaction between internal waves and the mean flow in the regime where the length and time scales of the waves are much shorter than those of the mean flow was first examined in detail by Bretherton [14] and Bretherton and Garrett [16] for small amplitude waves. Subsequently, this work was extended to finite amplitude waves, incorporating the perturbative effects of friction and compressibility, as well as the leading order effect of rotation, by Grimshaw [23]. Mean flows induced by internal waves propagating in a shear flow without rotation were also discussed by Grimshaw [24,25]. Relatively recent reviews of WMFI dynamics in the ocean appear in [26,37].

The energetics of WMFI have been traditionally studied by introducing a radiation stress tensor due to the waves and describing its role in the wave energy equation and the mean-flow equation. This approach follows ideas of Dewar [18] and Bretherton [15] for slowly varying linearized waves, and was further developed in the generalized Lagrangian mean (GLM) formulation of Andrews and McIntyre [7,8], who treated the Lagrangian displacement due to the waves as a *prescribed* field. Such radiation stress terms due to internal waves also emerge naturally in the momentum balance relations derived at each level of the present WMFI theory.

The particle-mechanics form of HP for the Lagrangian description of ideal fluid dynamics has long been known. See, e.g., [31,43] for reviews. However, the Lie algebraic significance of the Poisson bracket in the Hamiltonian formulation of Eulerian ideal fluid dynamics only emerged recently, first in the work of Arnold [9] for incompressible planar ideal flow. Holm and Kupershmidt [28,29], Holm et al. [30] and Marsden et al. [35] extended this result of Arnold to a wide variety of compressible ideal fluid and plasma theories, thereby making the Lie algebraic and group theoretic interpretations of the Lie–Poisson Hamiltonian formulation of Eulerian ideal fluid dynamics into a general principle. This general principle – that ideal Eulerian fluid dynamics is Lie–Poisson – has been found useful, for example, in establishing nonlinear stability theorems for ideal fluid and plasma equilibria by using a variant of the Liapunov method of stability analysis for Hamiltonian systems first convincingly demonstrated in this setting by Arnold [10,11]. The nonlinear stability method has since been applied to a great many hydrodynamic stability problems. For extensive reviews, see [2,32]. Shepherd [45] reviews recent applications of this method in geophysical fluid dynamics. An adaptation of the stability method has also been used by McIntyre and Shepherd

[36] to make a finite-amplitude extension of a conserved quantity found by Andrews [3] for small disturbances to a steady basic flow.

We emphasize that such nonlinear conserved quantities in Eulerian ideal fluid dynamics are not peculiarities of particular fluid theories and approximations. Rather, they are general properties of *all* Eulerian ideal fluid theories, including WMFI dynamics, because of the invariance of the Eulerian description under shifts of Lagrangian fluid parcels along streamlines of *steady* Eulerian flows. This invariance is obvious when HP – which only depends on the Eulerian variables – is formulated in terms of variations of Lagrangian fluid labels at fixed Eulerian position and time. In this formulation, the invariance of HP under such shifts of Lagrangian labels leads via Noether's theorem to an infinity of conservation laws for the potential vorticity integrals. In the corresponding Hamiltonian formulation of Eulerian fluid dynamics, these potential vorticity integrals are the "Casimirs" [32] of the Lie–Poisson bracket which Poisson-commute with all Eulerian variables. In the present case, averaging in HP produces additional conservation laws, since by definition averaging creates ignorable coordinates.

Sections 2 and 3 give two derivations of our leading order WMFI equations, and then discuss their conservation laws and Kelvin circulation theorem. The first derivation is based on the Euler equations, while the second is based on HP for the Euler equations. In Section 2.1, the fluid trajectory and pressure are decomposed into wave and mean-flow components. Substitution of this decomposition into the equations of motion for an ideal incompressible stratified flow in a rotating frame in Section 2.2, followed by projection of these equations onto their mean and rapidly varying components and truncation of their expansion in ϵ and α in Section 2.3, yields the leading order WMFI equations. These equations are summarized concisely in Eqs. (4.44)–(4.46). The WMFI equations derived in Section 2 by the direct expansion approach are rederived in Section 3 by substituting the *same* WMF decomposition into HP for Euler's equations, truncating its expansion in powers of ϵ and α, and then averaging over the rapid phase of the wave components before taking variations.

Various properties of the leading order WMFI equations are discussed in Section 4, including quasi- and semi-geostrophic balances (Section 4.1), initial conditions and balance relations (Section 4.2), as well as mean vorticity dynamics and the magnitude of wave effects relative to the beta effect (Section 4.3). The second derivation of the leading order WMFI equations using the phase-averaged HP places this theory into a framework – discussed in Section 4.4 – in which Noether's theorem may be applied systematically to derive the conservation laws of energy, potential vorticity, and wave action for the self-consistent WMFI theory from symmetries of its phase-averaged HP. Wave action conservation corresponds to the Eliassen–Palm relation in meteorology [3,5,6]. In Section 4.5 we discuss the solution procedure for determining the Eulerian mean pressure, which is necessary in order to close the WMFI theory. In Section 4.6 we derive the Kelvin circulation theorem for the leading order WMFI theory in the framework of the phase-averaged HP. This Kelvin circulation theorem is the basis for the Charney–Drazin "nonacceleration" theorem in WMFI which is well known in meteorology [17] – namely, that a steady wave produces no Lagrangian mean circulation, in the absence of dissipation, *provided* the wave pseudomomentum density has no curl. The leading order WMFI circulation theorem (4.40) extends to arbitrary fluid contours the circulation theorem of Grimshaw [23] for horizontal fluid contours – in the incompressible limit with the Boussinesq approximation, and when viscosity is absent. The wave pseudomomentum density at leading order in our WMFI theory is the product of wave action density and wave vector (the slow spatial gradient of the wave phase). The leading order total momentum balance relation (4.21) involves both fluid momentum density and wave pseudomomentum density in the momentum stress tensor (4.23). The latter terms comprise the radiation stresses invoked in previous WMFI theories. One may refer to, e.g., [4–6,15,18,36] for further discussions of Eliassen–Palm relations, internal wave radiation stress and the Charney–Drazin theorem.

In Section 5 the Hamiltonian formulation of WMFI dynamics is derived by Legendre-transforming the constrained and phase-averaged HP of Section 3. This Hamiltonian formulation yields a noncanonical Poisson bracket which is the sum of two Lie–Poisson Hamiltonian operators, one involving the mean-flow Eulerian fluid momentum density, the volume element of a fluid parcel and its buoyancy, and another involving the wave pseudomomentum density and the wave action density. The associated Lie–Poisson brackets are each dual to the semidirect-product Lie algebra of vector fields acting on functions and densities. (Here "dual" is meant in both the algebraic and metric senses, cf. [29,35].) Dual coordinates in the Lie–Poisson Hamiltonian operator for the mean-flow subsystem of WMFI are: Eulerian momentum density (dual to vector fields); the conserved incompressible volume element (dual to functions); and the stratified buoyancy (dual to densities). Dual coordinates in the Lie–Poisson Hamiltonian operator for the wave subsystem of WMFI are: wave pseudomomentum density (dual to vector fields) and the conserved wave action density (dual to functions). This separation of the Hamiltonian structure into a sum of two semidirect-product Lie–Poisson Hamiltonian operators which operate on different subsets of the fluid variables provides two limiting cases: one case in which the wave field propagates in a prescribed mean flow; and another in which the wave variables are prescribed and act as forces on the evolving mean flow of the fluid. The self-consistent coupling between these two subsystems occurs through the Hamiltonian and its constraints. Consequently, WMFI dynamics may be regarded as mean hydrodynamic motion coupled energetically and through constraints to the wave degree of freedom. This coupling has richer implications than a mere thermodynamic coupling, as would occur if, say, another degree of freedom were introduced into the fluid's equation of state, since neither the wave action density nor the wave pseudomomentum which transports it are "frozen" into the fluid's mean motion, as the wave propagates through the fluid. In fact, the semidirect-product Lie–Poisson Hamiltonian structure we find for the combined WMFI system is formally identical to that found for the two-fluid theory of superfluid He^4 by Holm and Kupershmidt [28]. Thus, wave mean-flow theory and the two-fluid theory of superfluid He^4 are analogous, in that both theories are Lie–Poisson two-fluid models whose order parameter is an S^1 phase. (The wave vector in WMFI theory is analogous to the (curl-free) superfluid velocity.) However, there is an important physical difference between the two theories: unlike the superfluid momentum density in the He^4 two-fluid model, the wave pseudomomentum density of the two-fluid WMFI system transports no mass or volume; it only transports wave action. Consequently, the two mathematically analogous theories have quite different physical behaviors.

At the end of Section 5, equilibrium solutions of the WMFI equations are characterized as critical points of the sum of the WMFI Hamiltonian, the integrated wave action density and the Casimir functions of the Lie–Poisson bracket. These Casimir functions are integrals of arbitrary functions of the locally conserved potential vorticity for WMFI dynamics. Establishing the WMFI equilibria as critical points of the sum of these conserved quantities places the theory into the framework of Hamiltonian nonlinear stability analysis, [32], whose application to these equations will be discussed elsewhere [22].

In Section 6 the computation of the averaged action and the associated WMFI equations is carried beyond the leading order, to higher powers in the asymptotic expansion in the adiabatic parameter ϵ and the wave amplitude α. The resulting expressions for the kinetic and potential energies in the averaged action are exact within the assumed decomposition; the only truncation occurs in the incompressibility constraint, at order $O(\alpha^4)$ in the averaged action, which affects the equations resulting from HP asymptotics at order $O(\alpha^4\epsilon)$. At this order, the Lagrangian mean flow is not quite incompressible; instead, a slight compressibility of order $O(\alpha^2\epsilon^2)$ arises from the wave motion. In this section, Noether's theorem is used to derive the conservation laws for WMFI energy, potential vorticity, and wave action to order $O(\alpha^4\epsilon)$. The Hamiltonian formulation of WMFI dynamics at this order again separates into wave and fluid subsystems [22] and the total momentum balance relation (6.21) involves both fluid and wave (radiation) terms in the higher-order momentum stress tensor (6.23). When the mean flow is prescribed, our higher-order WMFI equations reduce to equations analogous to those for WKB linear wave train stability analysis discussed by Lifschitz [33]. However, unlike the equations in [33], this restriction of the higher-order WMFI equations does not reduce to ordinary differential equations along fluid characteristics. When, instead of prescribing the mean flow, the wave displacement is prescribed, these higher-order WMFI equations reduce to the GLM equations [7] for the mean flow. In the process of considering this last restriction of the higher-order WMFI theory, we incidentally establish the existence of variational and Hamiltonian formulations of the GLM theory. In Appendix A, we discuss the properties of the WMFI theory at an intermediate order, when terms of order $O(\alpha^2\epsilon^2)$ are neglected in HP, as well as terms of order $O(\alpha^4)$. At this intermediate order, an additional phase symmetry exists in HP, whose associated conservation law implies preservation on fluid parcels of a certain wave polarization quantity analogous to wave angular momentum.

The main results of this paper are the derivations, Hamiltonian formulations, circulation theorems and conservation laws of self-consistent WMFI equations at various orders in α and ϵ, which are summarized in Sections 4.7 and 6.1, and in Appendix A. Other aspects of these equations will be discussed elsewhere.

2. Derivation of the leading order WMFI equations by the direct expansion approach

2.1. Decomposition of the fluid-parcel trajectory and pressure

In the Lagrangian description, Euler's equations for incompressible stratified fluid motion in a rotating frame follow from HP, $\delta\mathcal{L} = 0$, with the action $\mathcal{L}$ given by

$$\mathcal{L} = \int \mathrm{d}t \int \mathrm{d}^3L \left[\tfrac{1}{2}|\dot{\boldsymbol{X}}|^2 - \rho_1(L^A) g Z(L^A, t) + \dot{\boldsymbol{X}} \cdot \boldsymbol{R}(\boldsymbol{X}(L^A, t)) + p \left(\det\left(\frac{\partial X^i}{\partial L^A} \right) - 1 \right) \right], \tag{2.1}$$

where $X^i(L^A, t)$ is the position at time t of the fluid parcel located initially at position $\delta^i_A L^A$, $A = 1, 2, 3$. Thus, with a slight abuse of notation we may write

$$\boldsymbol{X}(L^A, 0) = L^A \quad \text{and} \quad \dot{\boldsymbol{X}}(L^A, t) = \left.\frac{\partial \boldsymbol{X}(L^A, t)}{\partial t}\right|_{L^A}, \tag{2.2}$$

where $\dot{\boldsymbol{X}}(L^A, t)$ is the current velocity of the fluid parcel labeled by L^A. In HP the independent variations of $\mathcal{L}$ are taken in the fluid trajectory $\boldsymbol{X}$ and pressure p at constant Lagrangian label L^A and time t. The Coriolis parameter for the rotating frame (twice the rotation vector) is the quantity $2\boldsymbol{\Omega}(\boldsymbol{X}) = \nabla \times \boldsymbol{R}(\boldsymbol{X})$, where $\boldsymbol{R}(\boldsymbol{X})$ is the vector potential for the Cariolis parameter, which is taken to be independent of time. The buoyancy of the fluid is $\rho_1(L^A)$, and the gravitational acceleration is g. Variations of $\mathcal{L}$ with respect to $\boldsymbol{X}$ and p yield Euler's equations in the Lagrangian description.

We choose to decompose the fluid trajectory into the sum of a mean flow $\boldsymbol{x}$ and a small, order $O(\alpha)$, displacement due the presence of an internal wave field whose phase (oscillation rate) is rapid compared to the rate of change of the mean flow in space and time. Namely,

$$\boldsymbol{X} = \boldsymbol{x} + \alpha\boldsymbol{\xi}(\boldsymbol{x}, t) \tag{2.3}$$

with

$$\boldsymbol{x}(l^A, t) = l^A + \int_O^t \dot{\boldsymbol{x}}(\epsilon l^A, \epsilon t')\, \mathrm{d}t', \tag{2.4}$$

where α and ϵ are small, real, positive parameters, $\epsilon \ll 1$, $\alpha \ll 1$, l^A is the initial condition for $\boldsymbol{x}$ and $\dot{\boldsymbol{x}}(\epsilon l^A, \epsilon t) = \partial\boldsymbol{x}(l^A, t)/\partial t$. Here ϵ denotes the adiabatic parameter (ratio of space and time scales between the wave oscillations and background flow) and α denotes the ratio of the wave amplitude and its typical wavelength. Note that $\epsilon\boldsymbol{x}$ depends only on ϵl^A and ϵt and, conversely, ϵl^A depends only on $\epsilon\boldsymbol{x}$ and ϵt. So it is consistent to regard the dependence on the Lagrangian labels as slow when it appears as ϵl^A, and as rapid when it appears as l^A (without ϵ). Since the nonlinear

Euler equations for an incompressible stratified flow under gravity in a rotating frame support plane internal waves on a uniform steady background as exact solutions, we may take the displacement $\boldsymbol{\xi}$ in the decomposition (2.3) to be a wave train with slow modulations in amplitude and phase

$$\boldsymbol{\xi}(\boldsymbol{x},t) = \boldsymbol{a}(\epsilon\boldsymbol{x},\epsilon t)\exp\left[\mathrm{i}\frac{\phi(\epsilon\boldsymbol{x},\epsilon t)}{\epsilon}\right] + \boldsymbol{a}^*(\epsilon\boldsymbol{x},\epsilon t)\exp\left[-\mathrm{i}\frac{\phi(\epsilon\boldsymbol{x},\epsilon t)}{\epsilon}\right] \tag{2.5}$$

and define its frequency as

$$\omega(\epsilon\boldsymbol{x},\epsilon t) = -\frac{\partial}{\partial\epsilon t}\phi(\epsilon\boldsymbol{x},\epsilon t) = -\frac{\partial}{\partial t}\frac{\phi(\epsilon\boldsymbol{x},\epsilon t)}{\epsilon} \tag{2.6}$$

and wave vector as

$$\boldsymbol{k}(\epsilon\boldsymbol{x},\epsilon t) = \frac{\partial}{\partial\epsilon\boldsymbol{x}}\phi(\epsilon\boldsymbol{x},\epsilon t) = \frac{\partial}{\partial\boldsymbol{x}}\frac{\phi(\epsilon\boldsymbol{x},\epsilon t)}{\epsilon}. \tag{2.7}$$

One of our main objectives is to show how the WMF decomposition (2.3) can be preserved, to a certain order, under time evolution according to the Euler equations, when $\boldsymbol{\xi}$ in Eq. (2.5) describes the fluid displacement due to a train of slowly modulated internal waves. The conditions for this preservation are the WMFI equations. We remark further on this choice for $\boldsymbol{\xi}$ in Section 2.4 and show that in the limit $\epsilon \to 0$ the WMFI equations possess an exact solution in the form of a plane internal wave. We then show that this solution is also an exact solution of the nonlinear Euler equations in this limit.

The decomposition of the fluid trajectory (2.3) implies the following decomposition of the fluid velocity:

$$\dot{\boldsymbol{X}}(L^A(l^B),t) = \dot{\boldsymbol{x}}(\epsilon l^A,\epsilon t) + \alpha\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}, \tag{2.8}$$

where for any function f of t and $\boldsymbol{x}(l^A,t)$,

$$\frac{\mathrm{d}f(\boldsymbol{x},t)}{\mathrm{d}t} \equiv \left.\frac{\partial f(\boldsymbol{x}(l^A,t),t}{\partial t}\right|_{l^A} = \left.\frac{\partial f}{\partial t}\right|_{\boldsymbol{x}} + \left.\frac{\partial f}{\partial\boldsymbol{x}}\right|_t \cdot\dot{\boldsymbol{x}}(\epsilon l^A,\epsilon t). \tag{2.9}$$

Since the transformation between L^A and l^A does not involve the time, the derivative with respect to t at constant l^A is the same as the one taken at constant L^A. For later use we define the Lagrangian mean velocity $\bar{\boldsymbol{u}}_{\mathrm{L}}(\epsilon\boldsymbol{x},\epsilon t)$ to be the phase average of the total velocity following a fluid particle with Lagrangian label l^A, evaluated at its current Eulerian position and time. Thus,

$$\dot{\boldsymbol{x}}(\epsilon l^A(\epsilon\boldsymbol{x},\epsilon t),\epsilon t) = \overline{\dot{\boldsymbol{X}}(L^A(l^B(\boldsymbol{x},t)),t)} = \bar{\boldsymbol{u}}_{\mathrm{L}}(\epsilon\boldsymbol{x},\epsilon t), \tag{2.10}$$

where overbar denotes phase average.

We decompose the pressure into its slowly and rapidly varying components as follows:

$$p(\boldsymbol{X},t) = p(\boldsymbol{x}+\alpha\boldsymbol{\xi},t) = p_0(\boldsymbol{x}+\alpha\boldsymbol{\xi},t) + \sum_{j\geq 1}\alpha^j p_j(\boldsymbol{x}+\alpha\boldsymbol{\xi},t), \tag{2.11}$$

where p_0 denotes the slowly varying part of the total pressure function, evaluated at the current position of the fluid parcel with label l^A. Thus,

$$p_0(\boldsymbol{x},t) = \tilde{p}_0(l^A(\boldsymbol{x},t),\epsilon t) \tag{2.12}$$

for a function $\tilde{p}_0$ and

$$p_j(\boldsymbol{x},t) = b_j(\epsilon\boldsymbol{x},\epsilon t)\exp\left[\mathrm{i}j\frac{\phi(\epsilon\boldsymbol{x},\epsilon t)}{\epsilon}\right] + b_j^*(\epsilon\boldsymbol{x},\epsilon t)\exp\left[-\mathrm{i}j\frac{\phi(\epsilon\boldsymbol{x},\epsilon t)}{\epsilon}\right] \quad \text{for } j\geq 1. \tag{2.13}$$

The pressure coefficients $p_j(\boldsymbol{x}, t)$ contain the rapidly fluctuating time dependence of the pressure. The pressure component p_0 is the Eulerian mean of the total pressure, $p_0(\boldsymbol{x}, t) = \overline{p(\boldsymbol{x}, t)}$, in which the phase average is taken at fixed Eulerian position $\boldsymbol{x}$. At Eulerian position $\boldsymbol{x}$, p_0 depends explicitly only slowly on time, but it may depend rapidly upon the Lagrangian labels $l^A(\boldsymbol{x}, t)$, $A = 1, 2, 3$. (In the Lagrangian description the fluid displacements $\boldsymbol{X}$ and $\boldsymbol{x}$ are expressed as functions of L^A and l^A, respectively, and of time.)

The pressure p can also be decomposed into the sum of its Lagrangian mean $\overline{p}_{\mathrm{L}}(\boldsymbol{x}, t) = \overline{p(\boldsymbol{x} + \alpha\boldsymbol{\xi}, t)}$ (in which the phase average is taken following a fluid trajectory) and its remaining Lagrangian rapid fluctuations. when imposing conditions on p that are required for the existence of internal waves, it turns out to be simpler and more natural to deal with the Eulerian decomposition (2.11)–(2.13) than with the decomposition into the sum of Lagrangian mean pressure and rapid fluctuations. We will return to this decomposition of p when we discuss prescribed fluctuations in Section 6.3.

2.2. Motion equations in the WMF decomposition

Without loss of generality, we assume that $\det(\partial L^A/\partial l^B) = 1$. Moreover, for simplicity in the calculation we set $\boldsymbol{R}(\boldsymbol{X}) = \boldsymbol{\Omega} \times \boldsymbol{X}$ with $\boldsymbol{\Omega} = \Omega\hat{\boldsymbol{z}}$ and Ω = const, so $\boldsymbol{\nabla} \times \boldsymbol{R} = 2\boldsymbol{\Omega}$. Allowing the spatial variation of $\boldsymbol{\Omega}$ is straightforward, but gives rise to complications in notation that we wish to avoid. In Section 4.3 we will consider the leading order effect of such spatial variations. With the decomposition (2.3) the action $\mathcal{L}$ of (2.1) becomes

$$\begin{aligned}\mathcal{L} = \int \mathrm{d}t \int \mathrm{d}^3 l \Bigg[& \frac{1}{2}\left|\dot{\boldsymbol{x}}(\epsilon l^A, \epsilon t) + \alpha\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}\right|^2 - \rho(l^A) g\Big(z(l^A, t) + \alpha\xi^3(\boldsymbol{x}(l^A, t), t)\Big) \\ & + \left(\dot{\boldsymbol{x}}(\epsilon l^A, \epsilon t) + \alpha\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}\right) \cdot \boldsymbol{R}\Big(\boldsymbol{x}(l^A, t) + \alpha\boldsymbol{\xi}(\boldsymbol{x}(l^A, t), t)\Big) \\ & + p\left(\det\left(\frac{\partial x^k}{\partial l^A}\right)\det\left[\delta^i_j + \alpha\left(\frac{\partial x^j}{\partial l^B}\right)^{-1}\frac{\partial \xi^i}{\partial l^B}\right] - 1\right)\Bigg],\end{aligned} \tag{2.14}$$

where $\rho(l^A) = \rho_{\mathrm{l}}(L^B(l^A))$. The independent variations of $\mathcal{L}$ are now taken in $\boldsymbol{x}$ and p at constant l^A and t.

Variation of $\mathcal{L}$ with respect to p yields the constraint

$$\det\left(\frac{\partial x^k}{\partial l^A}\right)\det\left[\delta^i_j + \alpha\left(\frac{\partial x^j}{\partial l^B}\right)^{-1}\frac{\partial \xi^i}{\partial l^B}\right] - 1 = 0 \tag{2.15}$$

and variation with respect to $\boldsymbol{x}$ yields the following motion equation:

$$\begin{aligned}& -\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{x}_m + \alpha\frac{\mathrm{d}\xi_m}{\mathrm{d}t}\right) - \alpha\frac{\partial \boldsymbol{\xi}}{\partial x^m}\cdot\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\boldsymbol{x}} + \alpha\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}\right) - \rho g\delta_{m3} - \alpha\rho g\frac{\partial \xi_3}{\partial x^m} \\ & + (\dot{\boldsymbol{x}} \times 2\boldsymbol{\Omega})_m + \alpha\frac{\partial \boldsymbol{r}}{\partial x^m}\cdot\dot{\boldsymbol{x}} - \alpha\frac{\partial r_m}{\partial x^i}\dot{x}^i - \alpha\frac{\partial r_m}{\partial t} - \alpha\frac{\partial \boldsymbol{\xi}}{\partial x^m}\cdot\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}t} + \alpha\frac{\partial \boldsymbol{R}}{\partial x^m}\cdot\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t} \\ & - \left(\frac{\partial x^m}{\partial l^A}\right)^{-1}\frac{\partial p}{\partial l^A} = 0.\end{aligned} \tag{2.16}$$

The quantity $\boldsymbol{r}$ is defined by $\boldsymbol{r} = \boldsymbol{\Omega} \times \boldsymbol{\xi}$. The order $\mathrm{O}(\alpha)$ terms in Eq. (2.16) are the linearized equations in the Lagrangian description for the stability of the mean flow $\dot{\boldsymbol{x}}$, when $\dot{\boldsymbol{x}}(l^A, t)$ is prescribed.

With the decompositions of the fluid trajectory and pressure in the forms (2.3) and (2.11), Eq. (2.15) and (2.16) contain terms that vary slowly in t, as well as those that are proportional to powers of the rapidly oscillating phase factor $\mathrm{e}^{\mathrm{i}\phi/\epsilon}$. To ensure that Euler's equations preserve the decompositions (2.3) and (2.11) through a certain order in α and ϵ, we must separately set to zero terms proportional to $e^{in\phi/\epsilon}$ where n successively takes on the values $0, \pm 1, \pm 2, \ldots$ Note that simply averaging Eq. (2.16) over the phase ϕ at constant l^A is inappropriate, as we shall discuss at the end of this section.

In order to use scaled variables ϵl^A and ϵt in the Lagrangian description, or $\epsilon \boldsymbol{x}$ and ϵt in the Eulerian description, we first divide Eq. (2.16) through by ϵ. (If Eq. (2.15) were written as a continuity equation, it too would need to be divided by ϵ for the leading order dynamical evolution in (2.16) to describe incompressible flow). Then, without specifying the relative magnitudes of α and ϵ, we consistently neglect terms of orders $O(\alpha^2\epsilon)$ and $O(\alpha^4)$ for the remainder of this section. This level of accuracy is sufficient to determine the leading effect of the waves on the mean flow, which occurs in our scaling at order $O(\alpha^2)$.

2.3. Projections of the motion onto its harmonic components

The following notation is used: The left-hand side of (2.15) is denoted by A, the left-hand side of (2.16) by ϵB, and the projection operator onto the basis vector $e^{in\phi/\epsilon}$ by P_n/α^n. Thus, the operator P_n projects out terms proportional to $e^{in\phi/\epsilon}$ and multiplies them by α^n. The projection P_0, in particular, picks out terms independent of the rapid oscillations in (2.15) and (2.16).

The equations resulting from $P_n A = 0$ and $P_n B = 0$ are written in the Eulerian description as follows. Setting $P_0 A = 0$ (cf. Eq. (2.15)) gives volume preservation to order $O(\alpha^2\epsilon^2)$,

$$D = 1 + O(\alpha^2\epsilon^2), \tag{2.17}$$

where $D \equiv \det(\partial l^A/\partial x^i) = \det(\partial(\epsilon l^A)/\partial \epsilon x^i)$. Setting $P_0 B = 0$ (cf. Eq. (2.16)) gives the WMFI motion equation

$$\begin{aligned}&\frac{\partial(\boldsymbol{m}/D)}{\partial \epsilon t} - \bar{\boldsymbol{u}}_L \times \left(\frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \frac{\boldsymbol{m}}{D}\right) + \frac{1}{\epsilon}\rho g \hat{\boldsymbol{z}} + \frac{\partial}{\partial \epsilon \boldsymbol{x}}\left(\frac{1}{2}|\bar{\boldsymbol{u}}_L|^2 + p_0\right) \\ &+ \alpha^2 \frac{\partial}{\partial \epsilon \boldsymbol{x}}\left[-\omega\frac{N}{D} + \tilde{\omega}^2|\boldsymbol{a}|^2 + \frac{1}{2}\frac{\partial^2 p_0}{\partial x^i \partial x^j}(a^i a^{*j} + a^{*j}a^j)\right] = O(\alpha^2\epsilon),\end{aligned} \tag{2.18}$$

where $\tilde{\omega}$ is the wave frequency Doppler-shifted by the Lagrangian mean motion,

$$\tilde{\omega} = \omega - \boldsymbol{k} \cdot \bar{\boldsymbol{u}}_L, \tag{2.19}$$

$\boldsymbol{m}$ is the "Eulerian momentum density",

$$\frac{\boldsymbol{m}}{D} = \bar{\boldsymbol{u}}_L - \alpha^2 \boldsymbol{k}\frac{N}{D} + (\boldsymbol{\Omega} \times \boldsymbol{x}) + O(\alpha^2\epsilon). \tag{2.20}$$

N is the wave action density,

$$\frac{N}{D} = 2|\boldsymbol{a}|^2\tilde{\omega} + 2i\boldsymbol{\Omega} \cdot (\boldsymbol{a} \times \boldsymbol{a}^*) + O(\epsilon), \tag{2.21}$$

and $\bar{\boldsymbol{u}}_L$ is the Lagrangian mean velocity,

$$\bar{\boldsymbol{u}}_L(\epsilon \boldsymbol{x}, \epsilon t) = \dot{\boldsymbol{x}}(\epsilon l^A(\epsilon \boldsymbol{x}, \epsilon t), \epsilon t), \tag{2.22}$$

cf. Eq. (2.10). The slow dependence on $\boldsymbol{x}$ and t of $\bar{\boldsymbol{u}}_L$ is inherited from the slow dependence on l^A and t of $\dot{\boldsymbol{x}}(\epsilon l^A, \epsilon t)$, cf. the remarks after Eq. (2.4).

At leading order, $O(1/\epsilon)$, the motion equation (2.18) establishes hydrostatic and geostrophic balances, namely

$$2\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_L + \rho g \hat{\boldsymbol{z}} + \frac{\partial p_0}{\partial \boldsymbol{x}} = 0. \tag{2.23}$$

In order to provide the restoring force for internal waves, the relative density (or, buoyancy) $\rho(l^A(\boldsymbol{x}, t))$ must have one derivative of order O(1) with respect to z. As we show in Section 4.5, this implies that the initial condition for ρ in terms of Lagrangian labels is expressed as

$$\rho = l^3 \rho'(\epsilon l^3) + \rho''(\epsilon l^A), \tag{2.24}$$

where ρ' and ρ'' are auxiliary functions. For conditions of hydrostatic balance, Eq. (2.23) assigns to the pressure p_0 two derivatives of order O(1) with respect to z. Depending on the initial conditions for $\dot{\boldsymbol{x}}$, the geostrophic balance in Eq. (2.23) may in addition require p_0 to have up to one derivative of order O(1) with respect to x and y. Since $\overline{\boldsymbol{u}}_{\mathrm{L}}$ by assumption depends only on slow spatial coordinates, however, the *mixed* second partial derivatives of p_0 must be of order $\mathrm{O}(\epsilon)$. Specifically, the function $\tilde{p}_0$ of Eq. (2.12) is of the form

$$\tilde{p}_0(l^A, \epsilon t) = (l^3)^2 p_0'(\epsilon l^3, \epsilon t) + l^B (p_0'')_B(\epsilon l^A, \epsilon t) + p_0'''(\epsilon l^A, \epsilon t), \tag{2.25}$$

where, as usual, we sum on the repeated index $B = 1, 2, 3$. The functions p' and p_0' in Eqs. (2.24) and (2.25) cannot depend on $l^a, a = 1, 2$, as we discuss in Section 4.5 Furthermore, in order for the hydrostatic and geostrophic balances in Eq. (2.23) to remain the leading order terms in the motion equation (2.18) for times up to order $\mathrm{O}(1/\epsilon)$, it is necessary that the Lagrangian mean vertical velocity be of order $\mathrm{O}(\epsilon)$, i.e.,

$$\overline{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{\boldsymbol{z}} = \epsilon w_1(\epsilon \boldsymbol{x}, \epsilon t), \tag{2.26}$$

as we show in full detail in Section 4.2. There are no constraints on the horizontal velocity components; so $\overline{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{\boldsymbol{x}}$ and $\overline{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{\boldsymbol{y}}$ can be of order O(1). We shall see in Section 4.5 that Eqs. (2.24)–(2.26) guarantee that the derivatives ρ and p_0 will have the required properties for times up to order $\mathrm{O}(1/\epsilon)$.

Setting $P_1 A = 0$ (cf. Eq. (2.15)) shows that volume preservation yields wave transversality, to order $\mathrm{O}(\epsilon)$,

$$\boldsymbol{k} \cdot \boldsymbol{a} = \mathrm{O}(\epsilon), \tag{2.27}$$

whereas the projection $P_1 B = 0$ (cf. Eq. (2.16)) gives at order $\mathrm{O}(\alpha^2/\epsilon)$ a linear equation for the vector wave amplitude, $\boldsymbol{a}$,

$$\tilde{\omega}^2 \boldsymbol{a} + 2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega} \times \boldsymbol{a}) - \mathrm{i} b_1 \boldsymbol{k} - \left(\boldsymbol{a} \cdot \frac{\partial}{\partial \boldsymbol{x}}\right) \frac{\partial p_0}{\partial \boldsymbol{x}} = \mathrm{O}(\epsilon). \tag{2.28}$$

As explained in the previous paragraph, the pressure p_0 defined in Eq. (2.11) has two derivatives of order O(1) only with respect to z. Therefore, for the last term in Eq. (2.28) we have

$$\left(\boldsymbol{a} \cdot \frac{\partial}{\partial \boldsymbol{x}}\right) \frac{\partial p_0}{\partial \boldsymbol{x}} = a^3 \frac{\partial^2 p_0}{\partial z^2} \hat{\boldsymbol{z}} + \mathrm{O}(\epsilon).$$

Nevertheless, for direct comparison with the higher-order calculation in Section 6, we prefer to leave this term in the form given in Eq. (2.28). Taking $-\mathrm{i}\boldsymbol{a}^* \cdot P_1 B + \mathrm{i}\boldsymbol{a} \cdot P_{-1} B = 0$ gives the wave action conservation equation

$$\frac{\partial N}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot \left(\overline{\boldsymbol{u}}_{\mathrm{L}} N + \mathrm{i} D(\boldsymbol{a}^* b - \boldsymbol{a} b^*)\right) = \mathrm{O}(\epsilon), \tag{2.29}$$

where N is given in (2.21). Note to obtain wave action conservation, it is essential to retain terms of order $\mathrm{O}(\alpha^2)$ in $P_1 B$. Apart from a scalar relation that determines b_2, which is discussed in Section 2.4, the projections $P_2 A = 0$ and $P_2 B = 0$ only contain terms of order $\mathrm{O}(\alpha^4 \epsilon)$ and, thus, represent wave–wave interactions, which one neglects in WMFI dynamics.

2.4. Remarks about the leading order WMFI equations

2.4.1. Lagrangian mean pressure

Decomposing the pressure p into its Lagrangian mean $\bar{p}_L$, plus rapidly oscillating pressure fluctuations still results in Eq. (2.28), provided $\bar{p}_L$ and the amplitudes of the pressure fluctuations all possess spatial derivatives of order O(1).

2.4.2. Method of averaging

The WMFI motion equation (2.18) cannot be viewed as a result of the application of the lowest order of the method of averaging to the motion equation (2.16) with a prescribed rapidly oscillating function $\boldsymbol{\xi}$, for two reasons. (For an overview of the method of averaging see, e.g., [34,44].) First, the motion equation (2.16) has not been explicitly solved for $\mathrm{d}\dot{\boldsymbol{x}}/\mathrm{d}t$; so it is not in the standard form for averaging. Averaging an equation in nonstandard form leads to equations whose solutions generally do not remain near the solutions of the unaveraged equation for times of order $\mathrm{O}(1/\epsilon)$, [34,44]. The difference between the Euler motion equation (2.16) and its standard form involves terms of order $\mathrm{O}(\alpha)$ which do not contain $\mathrm{d}\dot{\boldsymbol{x}}/\mathrm{d}t$. Second, even if Eq. (2.16) were put into the standard form for the method of averaging, the terms of order $\mathrm{O}(\alpha^2\epsilon)$ (since we divided by ϵ, in (2.18) these appear as order $\mathrm{O}(\alpha^2)$) which are retained in Eq. (2.18) and which are essential for describing the interaction between the waves and the mean flow, are smaller than the terms which would be obtained from the *second* order averaging of Eq. (2.16). Here the leading order is α^2. Therefore, the WMFI motion equation (2.18) is valid, only if the rapidly oscillating terms in Eq. (2.16) are forced to vanish separately, to order $\mathrm{O}(\alpha^2\epsilon^2, \alpha^4\epsilon)$. In other words, Eq. (2.18) is valid, provided Eqs. (2.27)–(2.29) also hold.

In contrast to the method of averaging, the method presented here is simply a decomposition of Eqs. (2.15) and (2.16) under the assumed form of the solutions (2.3)–(2.7) and (2.11)–(2.13). Its usefulness rests on the explicit appearance of small parameters α and ϵ and the consequent possibility of truncating various expansions of the equations and using perturbative methods to solve them. In order for the truncations to be valid, it is essential that the initial conditions are decomposed according to (2.3)–(2.7) and (2.11)–(2.13), and that these decompositions are preserved throughout the time evolution. Provided conditions (2.24)–(2.26) are satisfied initially, such self-consistent time evolution does indeed take place, as we shall discuss.

2.4.3. Plane internal-wave solutions

In the limit $\epsilon \to 0$, the WMFI equations possess an exact nonlinear solution in which $\boldsymbol{\xi}$ is a monochromatic plane wave of constant amplitude, frequency, and wave vector, and the latter two quantities satisfy the dispersion relation for internal waves. In this limit, Eqs. (2.17), (2.18), (2.27)–(2.29) imply

$$\begin{gathered} D - 1 = 0, \quad 2\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_L + \rho g\hat{\boldsymbol{z}} + \frac{\partial p_0}{\partial \boldsymbol{x}} = 0, \quad \boldsymbol{k} \cdot \boldsymbol{a} = 0, \\ \tilde{\omega}^2\boldsymbol{a} + 2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega} \times \boldsymbol{a}) - \mathrm{i}b_1\boldsymbol{k} - \left(\boldsymbol{a} \cdot \frac{\partial}{\partial \boldsymbol{x}}\right)\frac{\partial p_0}{\partial \boldsymbol{x}} = 0, \end{gathered} \tag{2.30}$$

where the right-hand sides vanish exactly since the neglected terms in the WMFI equations are all proportional to ϵ, and $\epsilon \to 0$. For $\bar{\boldsymbol{u}}_L$ we seek a solution of the form $\bar{\boldsymbol{u}}_L = (c_1, c_2, 0)$, where c_1 and c_2 are constants, and we have taken account of Eq. (2.26). Thus $\dot{\boldsymbol{x}} = (c_1, c_2, 0)$, $x_i = \delta_{iA}l^A + c_it$, $i = 1, 2$ and $x_3 = l_3$, cf. Eq. (2.4). From Eqs. (2.12), (2.24) and (2.25) it follows that

$$\rho = c_3 + zc_4 \quad p_0 = z^2c_5 + K_i(x_i - c_it) + K_3z, \quad i = 1, 2. \tag{2.31}$$

where c_3, c_4, c_5 and the vector $\boldsymbol{K}$ are all constant and of order O(1). Eqs. (2.30) then have a solution in which $\bar{\boldsymbol{u}}_\mathrm{L} = (c_1, c_2, 0)$, $\boldsymbol{a}, b, \boldsymbol{k}, \omega$, and $\tilde{\omega}$ are all of order O(1) and independent of both $\boldsymbol{x}$ and t, while ρ and p_0 are given by Eqs. (2.31), and $\tilde{\omega}$ and $\boldsymbol{k}$ are related by the dispersion relation for internal waves, cf. Eq. (3.24),

$$\tilde{\omega}^2 = \frac{(2\boldsymbol{\Omega}\cdot\boldsymbol{k})^2}{k^2} + k_3^2\frac{\partial^2 p_0}{\partial z^2}. \tag{2.32}$$

In the limit under discussion, all terms in (2.32) are independent of $\boldsymbol{x}$ and t. Thus the displacement is of the form in (2.5), $\boldsymbol{\xi} = \boldsymbol{a}\exp[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - wt)] + \boldsymbol{a}^*\exp[-\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x} - wt)]$.

As established previously, the WMFI equations are the projections of the full Euler equations onto the zeroth and first harmonics of $\mathrm{e}^{i\phi/\epsilon}$. Independently of ϵ, at order $\mathrm{O}(\alpha^4)$, a scalar relation, $b_2 = \tilde{\omega}^2\boldsymbol{a}\cdot\boldsymbol{a}$, arises from the projection onto the second harmonic, $\mathrm{e}^{2i\phi/\epsilon}$. All other harmonics higher than the first vanish indentically in the limit $\epsilon \to 0$. Therefore, as claimed earlier, the plane wave solution for $\boldsymbol{\xi}$ with constant amplitude, frequency, and wave vector satisfies the full nonlinear Euler equations.

In Section 3 we show that the WMFI equations (2.17)–(2.29) can be derived from the averaged HP $\delta\bar{\mathcal{L}} = 0$, where $\bar{\mathcal{L}}$ is the phase average of the action $\mathcal{L}$ in Eq. (2.14). We then examine the implications of this derivation. As we shall see, HP for equations (2.17)–(2.29) systematically implies energy conservation, Kelvin's circulation theorem, and convection of potential vorticity for the mean flow described by the WMFI equations.

3. Derivation of the leading order WMFI equations by HP asymptotics

3.1. Averaging in HP for ideal fluids

We return to HP (2.1) and pass to the Eulerian description, in which the action for Euler's equations is expressed as

$$\mathcal{L} = \int \mathrm{d}t \int \mathrm{d}^3X \left\{ \det\left(\frac{\partial L^A(\boldsymbol{X},t)}{\partial X^i}\right)\left[\frac{1}{2}|\boldsymbol{U}(\boldsymbol{X},t)|^2 - \rho_1(L^A(\boldsymbol{X},t))gZ \right.\right.$$
$$\left.\left. + \boldsymbol{U}(\boldsymbol{X},t)\cdot\boldsymbol{R}(\boldsymbol{X}(L^A,t))\right] - p\left(\det\left(\frac{\partial L^A}{\partial X^i}\right) - 1\right)\right\} \tag{3.1}$$

with the Eulerian fluid velocity defined by

$$\boldsymbol{U}(\boldsymbol{X},t) = \dot{\boldsymbol{X}}(L^A(\boldsymbol{X},t),t). \tag{3.2}$$

The independent variations are now to be taken in L^A and p at constant $\boldsymbol{X}$ and t. This transformation to the Eulerian description requires L^A to be an invertible function of $\boldsymbol{X}$ with derivatives $\partial L^A/\partial X^i$ which are continuous in both space and time. Invertibility is guaranteed by the constraint of volume preservation imposed by the Lagrange multiplier p. Continuity of the derivatives is an additional assumption.

We again choose the decomposition of the fluid trajectory given in Eqs. (2.3)–(2.7). In the Eulerian description this decomposition implies the following decomposition of the fluid velocity:

$$\boldsymbol{U}(\boldsymbol{X},t) = \boldsymbol{U}(\boldsymbol{x}+\alpha\boldsymbol{\xi},t) = \bar{\boldsymbol{u}}_\mathrm{L}(\epsilon\boldsymbol{x},\epsilon t) + \alpha\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\xi}, \tag{3.3}$$

where the Lagrangian mean velocity $\bar{\boldsymbol{u}}_\mathrm{L} \equiv \overline{\boldsymbol{U}(\boldsymbol{x}+\alpha\boldsymbol{\xi},t)}$ is given by Eq. (2.10). In addition, in the Eulerian description $\mathrm{d}/\mathrm{d}t$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t} = \frac{\partial}{\partial t} + \bar{\boldsymbol{u}}_\mathrm{L}\cdot\frac{\partial}{\partial \boldsymbol{x}}. \tag{3.4}$$

Note that Eq. (2.9) implies

$$\frac{\mathrm{d}l^A}{\mathrm{d}t} = 0. \tag{3.5}$$

Consequently, the components of the Lagrangian mean fluid velocity $\bar{\boldsymbol{u}}_{\mathrm{L}}^i$, $i = 1, 2, 3$, are expressible in terms of the partial derivatives of $l^A(\boldsymbol{x}, t)$ as

$$\bar{\boldsymbol{u}}_{\mathrm{L}}^i = -\left(D^{-1}\right)_A^i \frac{\partial l^A}{\partial t}, \quad i = 1, 2, 3, \tag{3.6}$$

with $D_i^A = \partial l^A/\partial x^i = \partial(\epsilon l^A)/\partial \epsilon x^i$ and the determinant $D = \det D_i^A$ (volume element) satisfies

$$\frac{\partial D}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot D\bar{\boldsymbol{u}}_{\mathrm{L}} = 0 \tag{3.7}$$

exactly. The decomposition of the pressure $p(\boldsymbol{X}, t)$ is given by Eqs. (2.11)–(2.13).

Eq. (3.5) can also be established from the relation

$$\left(\frac{\partial}{\partial t}\bigg|_{\boldsymbol{X}} + \boldsymbol{U} \cdot \frac{\partial}{\partial \boldsymbol{X}}\right) f(\boldsymbol{X}, t) = \left(\frac{\partial}{\partial t}\bigg|_{\boldsymbol{x}} + \bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \frac{\partial}{\partial \boldsymbol{x}}\right) f(\boldsymbol{x} + \alpha\boldsymbol{\xi}, t) \tag{3.8}$$

which holds for any f [7]. Using (3.8) on $\mathrm{d}L^A/\mathrm{d}t = 0$ and denoting

$$L^A\left(l^B(\boldsymbol{x}, t)\right) = L^A(\boldsymbol{x} + \alpha\boldsymbol{\xi}, t) \tag{3.9}$$

leads again to (3.5). (We have abused the notation in (3.9) by labeling the functions on both sides of the equation by L^A, without introducing additional notation.)

As before, we set $\det(\partial L^A/\partial l^B) = 1$ and $\boldsymbol{R}(\boldsymbol{X}) = \boldsymbol{\Omega} \times \boldsymbol{X}$, where $\boldsymbol{\Omega} = \Omega\hat{\boldsymbol{z}}$ and $\Omega = \text{const}$. After Taylor-expanding the pressure component $p_0(\boldsymbol{x} + \alpha\boldsymbol{\xi}, t)$ in the decomposition (2.11) in powers of α we average over the wave phase in the action (3.1) for Euler's equations while keeping l^A and ϵt fixed. Note that the averaging in this setting is a formal operation associated with the addition of a new degree of freedom to describe the wave [47]. Thus, averaging in itself does not entail any approximations. The approximations occur in the truncations of the expansions of the averaged action in the small parameters ϵ and α. The averaged action $\overline{\mathcal{L}}$ is

$$\begin{aligned}\overline{\mathcal{L}} = \int \mathrm{d}t \int \mathrm{d}^3x \Bigg\{ & D\bigg[\tfrac{1}{2}|\bar{\boldsymbol{u}}_{\mathrm{L}}|^2 + \alpha^2|\boldsymbol{a}|^2\tilde{\omega}^2 - \rho\left(l^A(\boldsymbol{x}, t)\right) gz + \bar{\boldsymbol{u}}_{\mathrm{L}} \cdot (\boldsymbol{\Omega} \times \boldsymbol{x}) \\ & + 2\mathrm{i}\alpha^2\tilde{\omega}\boldsymbol{\Omega} \cdot (\boldsymbol{a} \times \boldsymbol{a}^*)\bigg] + p_0(1 - D) + \mathrm{i}\alpha^2\epsilon p_0\boldsymbol{k} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}} \times (\boldsymbol{a} \times \boldsymbol{a}^*) \\ & + \alpha^2\left(b + a^m\frac{\partial p_0}{\partial x^m}\right)\left(\epsilon\frac{\partial a^{*j}}{\partial \epsilon x^j} - \mathrm{i}\boldsymbol{k} \cdot \boldsymbol{a}^*\right) + \alpha^2\left(b^* + a^{*m}\frac{\partial p_0}{\partial x^m}\right)\left(\epsilon\frac{\partial a^j}{\partial \epsilon x^j} + \mathrm{i}\boldsymbol{k} \cdot \boldsymbol{a}\right) \\ & + \alpha^2\left[\frac{1}{2}\frac{\partial^2 p_0}{\partial x^i \partial x^j}(a^i a^{*j} + a^j a^{*i}) + \mathrm{i}(b\boldsymbol{k} \cdot \boldsymbol{a}^* - b^*\boldsymbol{k} \cdot \boldsymbol{a})\right](1 - D) + \mathrm{O}(\alpha^2\epsilon, \alpha^4)\Bigg\}.\end{aligned} \tag{3.10}$$

Here $b(\epsilon x, \epsilon t)$ is $b_1(\epsilon\boldsymbol{x}, \epsilon t)$, with its index suppressed. Terms of $\mathrm{O}(\alpha^4)$ arise from the incompressibility constraint and are proportional to $(\boldsymbol{k} \cdot \boldsymbol{a})^{p_1}(\boldsymbol{k} \cdot \boldsymbol{a}^*)^{p_2}$, where the exponents p_1 and p_2 satisfy $p_1 + p_2 \geq 2$. When wave transversality is taken into account, cf. Eq. (3.19), then these terms give corrections of $\mathrm{O}(\alpha^4\epsilon)$ in the equations of motion. If, however, variations of $\overline{\mathcal{L}}$ higher than the first are computed, the neglected terms can be of order $\mathrm{O}(\alpha^4)$.

As discussed earlier, for internal waves to exist, the Eulerian mean pressure p_0 must possess two derivatives of order O(1) with respect to z and up to one derivative of order O(1) with respect to x and y. Likewise, the density

$\rho(l^A(\boldsymbol{x}, t))$ must have one derivative of order O(1) with respect to z. Thus, to find equations correct through order $O(\alpha^2)$, one should keep terms of order $O(\alpha^2)$ in the averaged action and *also* terms of order $O(\alpha^2\epsilon)$ when they include p_0 times a slow spatial derivative which could be shifted to p_0 by integration by parts in the course of taking variations. (So the neglected terms containing p_0 are of order $O(\alpha^2\epsilon^2)$.) The independent variations of $\overline{\mathcal{L}}$ are taken in ϵl^A, $\boldsymbol{a}$, ϕ, p_0, and b at constant $\boldsymbol{x}$ and t.

Integration by parts simplifies $\overline{\mathcal{L}}$ in expression (3.10) to

$$\begin{aligned}\overline{\mathcal{L}} = \int \mathrm{d}t \int \mathrm{d}^3x \Bigg\{ & D\Big[\frac{1}{2}|\overline{\boldsymbol{u}}_\mathrm{L}|^2 + \alpha^2|\boldsymbol{a}|^2\tilde{\omega}^2 - \rho(l(\boldsymbol{x},t))gz + \overline{\boldsymbol{u}}_\mathrm{L}\cdot(\boldsymbol{\Omega}\times\boldsymbol{x}) \\ & +2\mathrm{i}\alpha^2\tilde{\omega}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*)\Big] - \alpha^2\mathrm{i}D(b\boldsymbol{k}\cdot\boldsymbol{a}^* - b^*\boldsymbol{k}\cdot\boldsymbol{a}) \\ & +(1-D)\Big[p_0 + \frac{\alpha^2}{2}\frac{\partial^2 p_0}{\partial x^i \partial x^j}\left(a^i a^{*j} + a^j a^{*i}\right)\Big] \\ & +\alpha^2\epsilon\left(a^m \frac{\partial p_0}{\partial x^m}\frac{\partial a^{*j}}{\partial \epsilon x^j} + a^{*m}\frac{\partial p_0}{\partial x^m}\frac{\partial a^j}{\partial \epsilon x^j}\right) + \mathrm{O}\left(\alpha^2\epsilon, \alpha^4\right)\Bigg\},\end{aligned} \tag{3.11}$$

provided the quantity $\hat{\boldsymbol{n}}\cdot p_0\boldsymbol{k}\times(\boldsymbol{a}\times\boldsymbol{a}^*)$ vanishes on the boundary, as enforced by the Lagrange multipliers b and b^* in the interior of the flow. The term in $\overline{\mathcal{L}}$ involving b and b^* combines with last term in braces to give the phase-averaged work done by the fluid displacement due to the wave against the pressure force.

Another round of integration by parts, neglecting terms whose variations contribute only at higher orders and introducing additional notation, further simplifies the expression for $\overline{\mathcal{L}}$ and clarifies the coupling between its mean-flow and wave terms. Namely, $\overline{\mathcal{L}}$ splits into the sum of the average mean-flow action $\overline{\mathcal{L}}_\mathrm{MF}$ and the average wave action $\alpha^2\overline{\mathcal{L}}_\mathrm{W}$, given by

$$\begin{aligned}\overline{\mathcal{L}} &= \overline{\mathcal{L}}_\mathrm{MF} + \alpha^2\overline{\mathcal{L}}_\mathrm{W} \\ &= \int \mathrm{d}t \int \mathrm{d}^3x \left\{D\left[\tfrac{1}{2}|\overline{\boldsymbol{u}}_\mathrm{L}|^2 - \rho(l(\boldsymbol{x},t))gz + \overline{\boldsymbol{u}}_\mathrm{L}\cdot(\boldsymbol{\Omega}\times\boldsymbol{x})\right] + p_0(1-D) \right. \\ &\quad \left. +\alpha^2 F^{\mu *}D_{\mu\nu}F^\nu + \mathrm{O}(\alpha^2\epsilon, \alpha^4)\right\},\end{aligned} \tag{3.12}$$

where, strictly for notational convenience, we combine $\boldsymbol{a}$ and b into a "four-vector field" $F^\mu = (\boldsymbol{a}, b)$, with $\mu = 1, 2, 3, 4$, and define a Hermitian dispersion tensor $D_{\mu\nu} = D^*_{\nu\mu}$ given by

$$D_{ij} = D\tilde{\omega}^2\delta_{ij} - 2\mathrm{i}D\tilde{\omega}\epsilon_{ijk}\Omega_k - D\frac{\partial^2 p_0}{\partial x^i \partial x^j}, \qquad D_{4j} = \mathrm{i}Dk_j = -D_{j4}, \qquad D_{44} = 0. \tag{3.13}$$

It is clear from the decomposition of the WMFI action (3.12) that stationarity of $\overline{\mathcal{L}}_\mathrm{W}$ with respect to variations of the fields $F^{\mu *} = (\boldsymbol{a}^*, b^*)$ yields $D_{\mu\nu}F^\nu = 0$, whose solutions are the polarization eigendirections of the field F^μ, up to an overall complex constant, and whose solvability condition produces the dispersion relation for the waves. Likewise, stationarity of $\overline{\mathcal{L}}$ under variations with respect to the fluid variables produces equations for the mean flow, with order $O(\alpha^2)$ wave forcing which arises from the dependence of $D_{\mu\nu}$ on D and $\tilde{\omega} = \omega - \overline{\boldsymbol{u}}_\mathrm{L}\cdot\boldsymbol{k}$.

3.2. Equations arising from the averaged HP

Stationarity of $\overline{\mathcal{L}}$ under variations with respect to p_0 at fixed $\boldsymbol{x}$ and t implies volume preservation, namely, cf. Eq. (2.17)

$$D = 1 + \mathrm{O}(\alpha^2\epsilon^2, \alpha^4\epsilon), \quad \text{which implies} \quad \frac{\partial}{\partial \epsilon \boldsymbol{x}}\cdot\overline{\boldsymbol{u}}_\mathrm{L} = \mathrm{O}(\alpha^2\epsilon^2, \alpha^4\epsilon). \tag{3.14}$$

Likewise, stationarity under variations of $\overline{\mathcal{L}}$ in ϵl^A at fixed $\boldsymbol{x}$ and t yields the motion equation (2.18):

$$\frac{\partial(\boldsymbol{m}/D)}{\partial\epsilon t} - \overline{\boldsymbol{u}}_{\mathrm{L}} \times \left(\frac{\partial}{\partial\epsilon\boldsymbol{x}} \times \frac{\boldsymbol{m}}{D}\right) + \frac{1}{\epsilon}\rho g\hat{\boldsymbol{z}} + \frac{\partial}{\partial\epsilon\boldsymbol{x}}\left(\frac{1}{2}|\overline{\boldsymbol{u}}_{\mathrm{L}}|^2 + p_0\right)$$
$$+ \alpha^2\frac{\partial}{\partial\epsilon\boldsymbol{x}}\left[-\omega\frac{N}{D} + \tilde{\omega}^2|\boldsymbol{a}|^2 + \frac{1}{2}\frac{\partial^2 p_0}{\partial x^i\partial x^j}(a^i a^{*j} + a^{*i}a^j)\right] = \mathrm{O}(\alpha^2\epsilon) \tag{3.15}$$

with

$$\boldsymbol{m} = \frac{\delta\overline{\mathcal{L}}}{\delta\overline{\boldsymbol{u}}_{\mathrm{L}}} = D\overline{\boldsymbol{u}}_{\mathrm{L}} - \alpha^2\boldsymbol{k}N + D(\boldsymbol{\Omega} \times \boldsymbol{x}) + \mathrm{O}(\alpha^2\epsilon) \tag{3.16}$$

and

$$N = \frac{1}{\alpha^2}\frac{\delta\overline{\mathcal{L}}}{\delta\omega} = 2D|\mathrm{a}|^2\tilde{\omega} + 2\mathrm{i}D\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*) + \mathrm{O}(\epsilon) \tag{3.17}$$

as defined in Eqs. (2.20) and (2.21). Thus N is proportional to the momentum density canonically conjugate to the wave phase ϕ. Likewise, the Eulerian momentum density $\boldsymbol{m}$ is related to the momentum $\pi_A \equiv \delta\overline{\mathcal{L}}/\delta l^A{}_{,t}$ canonically conjugate to l^A by

$$m_i = -\pi_A D_i^A = -\pi_A\frac{\partial l^A}{\partial x^i} \quad \text{or} \quad \boldsymbol{m} = -\pi_A\nabla l^A, \tag{3.18}$$

where we have used the relation (3.6) in taking variations to calculate π_A.

Stationarity of $\overline{\mathcal{L}}$ under variations in b^* and $\boldsymbol{a}^*$ gives, respectively,

$$\boldsymbol{k}\cdot\boldsymbol{a} = \mathrm{O}(\epsilon) \tag{3.19}$$

and

$$\tilde{\omega}^2\mathrm{a} - 2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega}\times\boldsymbol{a}) - \mathrm{i}b\boldsymbol{k} - \left(\boldsymbol{a}\cdot\frac{\partial}{\partial\boldsymbol{x}}\right)\frac{\partial p_0}{\partial\boldsymbol{x}} = \mathrm{O}(\epsilon), \tag{3.20}$$

which recovers Eqs. (2.27) and (2.28).

As might be expected for a phase-averaged Lagrangian, $\overline{\mathcal{L}}$ is invariant under a shift in the origin of the phase. Thus, ϕ is an ignorable coordinate and stationarity of $\overline{\mathcal{L}}$ under variations in ϕ gives a conservation law for its canonically conjugate momentum:

$$\frac{\partial N}{\partial\epsilon t} + \frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot(\overline{\boldsymbol{u}}_{\mathrm{L}}N + \mathrm{i}D(\boldsymbol{a}^*b - \boldsymbol{a}b^*)) = \mathrm{O}(\epsilon). \tag{3.21}$$

The phase-averaged Lagrangian $\overline{\mathcal{L}}$ is also invariant under the restricted phase shift

$$\boldsymbol{a} \to \mathrm{e}^{\mathrm{i}\gamma_0}\boldsymbol{a}, \quad b \to \mathrm{e}^{\mathrm{i}\gamma_0}b, \quad \gamma_0 = \text{const} \tag{3.22}$$

with all other variables left unchanged. However, this invariance is trivial at this order – its conserved density is a pure gauge.

In taking these variations we have set

$$(\hat{\boldsymbol{n}}\cdot\delta\boldsymbol{a}^*)\left(\boldsymbol{a}\cdot\frac{\partial p_0}{\partial\boldsymbol{x}}\right) = 0 \quad \text{and} \quad \delta\phi\hat{\boldsymbol{n}}\cdot(\overline{\boldsymbol{u}}_{\mathrm{L}}N + \mathrm{i}D(\boldsymbol{a}^*b - \boldsymbol{a}b^*)) = 0, \tag{3.23}$$

on the boundary, when integrating by parts. So the displacement of a fluid element and the flux of wave action density (which is N times the group velocity of the wave) are taken to be tangential to the boundary.

Remarks.

(1) Other boundary conditions could be imposed by using, say, the standard technique of adding a null Lagrangian to L. (A null Lagrangian is the space and time integral of a total divergence, whose only contribution in HP appears at the boundary. See, e.g., [19] for more discussion.) However, this approach is not pursued here.

(2) Eq. (3.21) can be viewed either as the result of varying $\overline{\mathcal{L}}$ with respect to ϕ, or as the conservation law arising via Noether's theorem from the invariance of $\overline{\mathcal{L}}$ under translations in ϕ.

(3) As stated after Eq. (2.28), the last term in Eq. (3.20) can also be written as $a^3\partial^2 p_0/\partial z^2\hat{z} + \mathrm{O}(\epsilon)$. However, the form in (3.20) is preferred for later comparison with higher-order calculations, cf. Eq. (6.12).

(4) Wave transversality (3.19) holds to order $\mathrm{O}(\epsilon)$, so Eq. (3.20) may be viewed as an eigenvalue equation for the two components of $\boldsymbol{a}$ perpendicular to $\boldsymbol{k}$, while the third component of the equation determines b. The solvability condition for the eigenvalue problem is the dispersion relation, which may be found, e.g., by taking $\boldsymbol{k} \times (\boldsymbol{k} \times (3.20))$ and $\boldsymbol{a} \cdot (3.20)$. Namely,

$$\tilde{\omega}^2 = \frac{(2\boldsymbol{\Omega}\cdot\boldsymbol{k})^2}{k^2} + \left(\delta^{jl} - \frac{k^j k^l}{k^2}\right)\frac{\partial^2 p_0}{\partial x^j \partial x^l} + \mathrm{O}(\epsilon). \tag{3.24}$$

Thus the Doppler-shifted frequency $\tilde{\omega}$ is independent of the magnitude of the wave vector $\boldsymbol{k}$ and depends on the mean flow through p_0. If the quantity $\partial^2 p_0/(\partial x^j \partial x^l)$ initially depends only on $\epsilon\boldsymbol{x}$, ϵt then it will remain so, as we will see in Section 4.5, provided the buoyancy ρ satisfies the initial condition (2.24), which supports internal waves.

(5) Under conditions of hydrostatic balance and when $\overline{\boldsymbol{u}}_{\mathrm{L}} = 0$, Eq. (3.24) reduces to the well-known dispersion relation for linear internal waves. Note that the initial conditions (2.24) and (2.25) on the buoyancy ρ and the Eulerian mean pressure p_0, and the requirement (2.26) on the vertical component of $\overline{\boldsymbol{u}}_{\mathrm{L}}$ imply that $\tilde{\omega}^2$ is order $\mathrm{O}(1)$, so we are not discussing critical layers (for which $\tilde{\omega} = 0$).

4. Properties of the leading order WMFI equations

4.1. *Quasi- and semi-geostrophic balances*

Upon manipulating the equations developed in the previous section, the motion equation (3.15) can be written in terms of $\overline{\boldsymbol{u}}_{\mathrm{L}}$ as

$$\left(\frac{\partial}{\partial \epsilon t} + \overline{\boldsymbol{u}}_{\mathrm{L}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}}\right)\overline{\boldsymbol{u}}_{\mathrm{L}} - \frac{1}{\epsilon}\left(\overline{\boldsymbol{u}}_{\mathrm{L}} \times 2\boldsymbol{\Omega} - \rho g \hat{\boldsymbol{z}} - \frac{\partial p_0}{\partial \boldsymbol{x}}\right)$$
$$= -\alpha^2\left[\frac{N}{D}\frac{\partial \tilde{\omega}}{\partial \epsilon \boldsymbol{x}} + \mathrm{i}\boldsymbol{k}\left(\frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot (\boldsymbol{a}^* b - \boldsymbol{a} b^*)\right)\right] + \mathrm{O}(\alpha^2\epsilon). \tag{4.1}$$

The leading order WMFI equations (3.14), (3.19)–(3.21) and (4.1) (which are collected in Section 4.7) describe dynamics on an invariant "slow manifold", which is nearly hydrostatically and geostrophically balanced. This is the sense in which the wave-mean flow equations comprise a reduction of the original Euler equations: Additional variables and equations are required, but these equations have solutions which vary slowly in space and time.

The motion equation (4.1) has terms of three different orders: $1/\epsilon$, 1, and α^2. At leading order (4.1) reduces to

$$2\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_{\rm L} + \rho g \hat{\boldsymbol{z}} + \frac{\partial p_0}{\partial \boldsymbol{x}} = 0, \tag{4.2}$$

which enforces the hydrostatic and geostrophic balances, as discussed before. Inclusion of terms of order O(1) in (4.1) yields the equations for the mean motion, $\bar{\boldsymbol{u}}_{\rm L}$, varying slowly in time and with no wave effects. This is also the order at which the nonlinearity occurs. Terms of order $O(\alpha^2)$ in (4.1) incorporate the "rectified" effects of rapidly fluctuating waves on the slowly varying motion. Thus the mean flow is hydrostatic and geostrophic to order O(1) and the adiabatic parameter ϵ formally plays the role of the Rossby number for the mean flow. The familiar quasi-geostrophic and semi-geostrophic relations may now be derived [40], with Rossby number replaced by the adiabatic parameter ϵ and with appropriate revisions to account for the effects of wave activity, which enter at order $O(\alpha^2)$.

4.2. *Initial conditions and balance relations*

The leading order hydrostatic and geostrophic balances of these equations hold, provided conditions (2.24)–(2.26) are satisfied, where the Lagrangian labels l^A (the initial coordinates) satisfy the advection law (3.5). Conditions (2.24)–(2.26) restrict the allowed initial conditions for the wave-current flow. Under these restrictions, the evolution remains for times of order $O(1/\epsilon)$ hydrostatically and geostrophically balanced to order O(1) and the rectified effects of the wave motion remain weak at order $O(\alpha^2)$.

We now demonstrate relation (2.26). Assuming that $\bar{\boldsymbol{u}}_{\rm L} \cdot \hat{\boldsymbol{z}}$ is of the form

$$\bar{\boldsymbol{u}}_{\rm L} \cdot \hat{\boldsymbol{z}} = w_0(\epsilon \boldsymbol{x}, \epsilon t) + \epsilon w_1(\epsilon \boldsymbol{x}, \epsilon t), \tag{4.3}$$

we show, in three steps, that w_0 must vanish. The first step establishes the independence of w_0 of transverse spatial coordinates, the second its independence of ϵz, and the third one invokes the boundary conditions to show that w_0 must be absent.

For the geostrophic balance in Eq. (4.1) to hold at order $O(1/\epsilon)$ for times up to order $O(1/\epsilon)$, and for $\bar{\boldsymbol{u}}_{\rm L}$ to depend only on slow spatial variables as assumed, p_0 must have exactly one derivative of order O(1) with respect to x and y. Taking this into account, as well as the form of p_0 given by Eqs. (2.12) and (2.25), it follows that for times up to order $O(1/\epsilon)$ the following relation holds :

$$\frac{\partial l^3}{\partial x^i} \le O(\epsilon), \quad i = 1, 2. \tag{4.4}$$

Therefore, the (3,1) and (3,2) elements of the Jacobian matrix D_i^A must be of order smaller than or equal to ϵ. Inversion of the matrix D_i^A then shows that for times up to order $O(1/\epsilon)$

$$\frac{\partial z}{\partial l^a} \le O(\epsilon), \quad a = 1, 2. \tag{4.5}$$

Use of Eq. (2.4) together with the Mean Value Theorem now implies that to order O(1) the z component of the mean velocity in the Lagrangian description, $\dot{x}^3(\epsilon l^A, \epsilon t)$, must be independent of ϵl^a, $a = 1, 2$. Together with relations (2.10) and (4.4), this implies that w_0 must be independent of ϵx and ϵy. That is, w_0 can only be a function of ϵz and ϵt.

To carry out the second step, we multiply both sides of Eq. (4.1) by ϵ and neglect terms of order $O(\alpha^2 \epsilon)$, with the result

$$\epsilon\left(\frac{\partial}{\partial \epsilon t} + \bar{\boldsymbol{u}}_{\rm L} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}}\right)\bar{\boldsymbol{u}}_{\rm L} + 2\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_{\rm L} + \rho g \hat{\boldsymbol{z}} + \frac{\partial p_0}{\partial \boldsymbol{x}} = O(\alpha^2 \epsilon). \tag{4.6}$$

We have neglected the terms that describe wave-current interactions, which enter at order $O(\alpha^2)$ in Eq. (4.1). The present argument only makes use of terms of the two lowest orders in Eq. (4.1), so the wave-current interactions do not play a role. We act on Eq. (4.6) with the operator $\epsilon(\partial/\partial\epsilon t + \bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \partial/\partial\epsilon\boldsymbol{x})$, use the relation $\boldsymbol{\Omega} = \Omega\hat{z}$ and the requirements that p_0 have two derivatives of order O(1) with respect to z and one derivative of order O(1) with respect to x and y, and perform one iteration in ϵ to solve for the horizontal component of $\bar{\boldsymbol{u}}_{\mathrm{L}}$. The result is

$$(\bar{\boldsymbol{u}}_{\mathrm{L}})_{\mathrm{H}} = \frac{1}{2\Omega}\left(\hat{z} \times \frac{\partial p_0}{\partial \boldsymbol{x}}\right) - \frac{\epsilon}{4\Omega^2}\left[\frac{\partial}{\partial\epsilon t} + \frac{1}{2\Omega}\left(\hat{z} \times \frac{\partial p_0}{\partial \boldsymbol{x}}\right) \cdot \frac{\partial}{\partial\epsilon\boldsymbol{x}_{\mathrm{H}}} + w_0\frac{\partial}{\partial\epsilon z}\right]\frac{\partial p_0}{\partial \boldsymbol{x}_{\mathrm{H}}} + O(\alpha^2\epsilon, \epsilon^2), \tag{4.7}$$

where the subscript H denotes the horiontal component: $(\bar{\boldsymbol{u}}_{\mathrm{L}})_{\mathrm{H}}$ is defined by $(\bar{\boldsymbol{u}}_{\mathrm{L}})_{\mathrm{H}} = \bar{\boldsymbol{u}}_{\mathrm{L}} - \hat{z}(\bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{z})$, and $\boldsymbol{x}_{\mathrm{H}}$ by $\boldsymbol{x}_{\mathrm{H}} = \boldsymbol{x} - \hat{z}(\boldsymbol{x} \cdot \hat{z})$. Taking the slow divergence of Eq. (4.7) then leads to the relation

$$\begin{aligned}\frac{\partial}{\partial\epsilon\boldsymbol{x}_{\mathrm{H}}} \cdot (\bar{\boldsymbol{u}})_{\mathrm{H}} = &-\frac{\epsilon}{4\Omega^2}\left[\frac{\partial}{\partial\epsilon t} + \frac{1}{2\Omega}\left(\hat{z} \times \frac{\partial p_0}{\partial \boldsymbol{x}}\right) \cdot \frac{\partial}{\partial\epsilon\boldsymbol{x}_{\mathrm{H}}}\right]\frac{\partial^2 p_0}{\partial\epsilon x^i_{\mathrm{H}}\partial x_{\mathrm{H}i}} \\ &-\frac{\epsilon}{4\Omega^2}\left[w_0\frac{\partial}{\partial\epsilon z}\right]\frac{\partial^2 p_0}{\partial\epsilon x^i_{\mathrm{H}}\partial x_{\mathrm{H}i}} + O(\alpha^2\epsilon, \epsilon^2),\end{aligned} \tag{4.8}$$

which by incompressibility, Eq. (3.14), implies

$$\partial w_0/\partial\epsilon z = O(\epsilon). \tag{4.9}$$

Therefore, in $\bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{z}$ all the dependence on ϵz is in w_1, so that w_0 can only be a function of ϵt.

The third step in showing that w_0 vanishes is straightforward. For a smooth boundary there exists a point where $\bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{z}$, and hence w_0, vanishes for all times, because of boundary conditions. Since w_0 at this point has been shown to be independent of spatial coordinates, it must therefore vanish throughout the enclosed domain for all times, which completes the demonstration that w_0 is indeed absent, so that $\bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{z} = O(\epsilon)$.

4.3. Lagrangian and Eulerian mean vorticity dynamics in WMFI

The curl of the motion equation in (4.1) yields the Lagrangian mean vorticity equation

$$\begin{aligned}&\epsilon\frac{\partial\bar{\boldsymbol{\omega}}_{\mathrm{L}}}{\partial\epsilon t} - \frac{\partial}{\partial\epsilon\boldsymbol{x}} \times (\bar{\boldsymbol{u}}_{\mathrm{L}} \times (\epsilon\bar{\boldsymbol{\omega}}_{\mathrm{L}} + 2\boldsymbol{\Omega})) + \frac{g}{\epsilon}\frac{\partial\rho}{\partial\boldsymbol{x}} \times \frac{\partial z}{\partial\boldsymbol{x}} \\ &= -\alpha^2\epsilon\left[\frac{\partial(N/D)}{\partial\epsilon\boldsymbol{x}} \times \frac{\partial\tilde{\omega}}{\partial\epsilon\boldsymbol{x}} + \mathrm{i}\frac{\partial}{\partial\epsilon\boldsymbol{x}}\left(\frac{\partial}{\partial\epsilon\boldsymbol{x}} \cdot (\boldsymbol{a}^*b - \boldsymbol{a}b^*)\right) \times \boldsymbol{k}\right] + O(\alpha^2\epsilon^2),\end{aligned} \tag{4.10}$$

where the Lagrangian mean vorticity $\bar{\boldsymbol{\omega}}_{\mathrm{L}}$ is given by

$$\bar{\boldsymbol{\omega}}_{\mathrm{L}} = \frac{\partial}{\partial\epsilon\boldsymbol{x}} \times \bar{\boldsymbol{u}}_{\mathrm{L}}. \tag{4.11}$$

Thus, wave activity may create Lagrangian mean fluid circulation (or Lagrangian mean vorticity) at a rate of order $O(\alpha^2)$ in slow time, when the gradients of the wave action density N and the Doppler-shifted wave frequency $\tilde{\omega}$ are not aligned. Likewise for $\boldsymbol{k}$ and the gradient of divergence of $\mathrm{i}(\boldsymbol{a}^*b - \boldsymbol{a}b^*)$ with respect to slow space. The quantity $\mathrm{i}D(\boldsymbol{a}^*b - \boldsymbol{a}b^*)$ is the flux of wave action which is not convected by the fluid, cf. the wave action conservation equation (3.21). So $\mathrm{i}D(\boldsymbol{a}^*b - \boldsymbol{a}b^*)/N$ is interpretable as $\boldsymbol{v}_{\mathrm{g}} - \bar{\boldsymbol{u}}_{\mathrm{L}}$, the group velocity of the waves relative to the Lagrangian mean velocity.

In order to examine further the physical significance of leading order terms in the motion equation (4.1), it is useful to introduce the Eulerian mean velocity $\bar{\boldsymbol{u}}_{\mathrm{E}}$. This quantity is defined to be the average of the exact velocity $\boldsymbol{U}$ of Eq. (3.2) at a fixed Eulerian position $\boldsymbol{x}$,

$$\bar{\boldsymbol{u}}_{\mathrm{E}} = \overline{U(\boldsymbol{x}, t)}. \tag{4.12}$$

The difference between $\bar{\boldsymbol{u}}_{\mathrm{L}}$ and $\bar{\boldsymbol{u}}_{\mathrm{E}}$ is the Stokes mean drift velocity $\bar{\boldsymbol{u}}_{\mathrm{S}}$ [7], which may be calculated in WMFI scaling as

$$\begin{aligned}
\bar{\boldsymbol{u}}_{\mathrm{S}} = \bar{\boldsymbol{u}}_{\mathrm{L}} - \bar{\boldsymbol{u}}_{\mathrm{E}} &= \overline{U(\boldsymbol{x} + \alpha\boldsymbol{\xi}, t)} - \overline{U(\boldsymbol{x}, t)} \\
&= \alpha^2 \left(\overline{\xi^i \frac{\partial}{\partial x^i} \frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}} \right) - \alpha^2 \overline{\xi^i \frac{\partial \xi^j}{\partial x^i}} \frac{\partial \bar{\boldsymbol{u}}_{\mathrm{L}}}{\partial x^j} - \frac{\alpha^2}{2} \overline{\xi^i \xi^j} \frac{\partial^2 \bar{\boldsymbol{u}}_{\mathrm{L}}}{\partial x^i \partial x^j} + \mathrm{O}(\alpha^4) \\
&= \alpha^2 \left(\overline{\xi^i \frac{\partial}{\partial x^i} \frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}} \right) + \mathrm{O}(\alpha^2\epsilon^2, \alpha^4) = \alpha^2 \epsilon \frac{\partial}{\partial \epsilon x^i} \left(\overline{\xi^i \frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t}} \right) + \mathrm{O}(\alpha^2\epsilon^2, \alpha^4),
\end{aligned} \tag{4.13}$$

where in the last step we have invoked $\partial \xi^i / \partial x^i = 0$, which appears at order $\mathrm{O}(\alpha)$ (the trace) in Eq. (2.15). Upon using the explicit form for $\boldsymbol{\xi}$ given by Eq. (2.5), the expression for $\bar{\boldsymbol{u}}_{\mathrm{S}}$ becomes

$$\bar{\boldsymbol{u}}_{\mathrm{S}} = \alpha^2 \epsilon \mathrm{i} \frac{\partial}{\partial \epsilon x^i} \left[\tilde{\omega} \left(a^i \boldsymbol{a}^* - a^{*i} \boldsymbol{a} \right) \right] - \mathrm{O}(\alpha^2\epsilon^2, \alpha^4). \tag{4.14}$$

Thus, $\bar{\boldsymbol{u}}_{\mathrm{S}}$ appears as an order $\mathrm{O}(\alpha^2\epsilon)$ quantity in both Lagrangian and Eulerian mean vorticity dynamics in WMFI.

Remark on the beta effect. Suppose small spatial variations of the rotation vector $\boldsymbol{\Omega}$ are considered for a moment, by setting

$$\boldsymbol{\Omega} = \Omega \hat{\boldsymbol{z}} + \beta \boldsymbol{\Omega}_1(\boldsymbol{x}), \tag{4.15}$$

where $\Omega = \mathrm{const}$ and $\beta \ll 1$. With the definition of the Eulerian mean vorticity $\bar{\boldsymbol{\omega}}_{\mathrm{E}}$

$$\bar{\boldsymbol{\omega}}_{\mathrm{E}} = \frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \bar{\boldsymbol{u}}_{\mathrm{E}}, \tag{4.16}$$

we take the curl of Eq. (4.1) and express $\bar{\boldsymbol{u}}_{\mathrm{L}}$ as $\bar{\boldsymbol{u}}_{\mathrm{L}} = \bar{\boldsymbol{u}}_{\mathrm{E}} + \bar{\boldsymbol{u}}_{\mathrm{S}}$ to get the equation for $\bar{\boldsymbol{\omega}}_{\mathrm{E}}$,

$$\begin{aligned}
&\frac{\partial \bar{\boldsymbol{\omega}}_{\mathrm{E}}}{\partial \epsilon t} + \bar{\boldsymbol{u}}_{\mathrm{E}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}} \bar{\boldsymbol{\omega}}_{\mathrm{E}} - \bar{\boldsymbol{\omega}}_{\mathrm{E}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}} \bar{\boldsymbol{u}}_{\mathrm{E}} - \frac{1}{\epsilon} \frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \left(\bar{\boldsymbol{u}}_{\mathrm{E}} \times 2\boldsymbol{\Omega} - \rho g \hat{\boldsymbol{z}} - \frac{\partial p_0}{\partial \boldsymbol{x}} \right) \\
&\quad + \frac{\beta}{\epsilon} \bar{\boldsymbol{u}}_{\mathrm{E}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}} 2\boldsymbol{\Omega}_1 - \frac{\beta}{\epsilon} 2\boldsymbol{\Omega}_1 \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}} \bar{\boldsymbol{u}}_{\mathrm{E}} = \mathrm{O}(\alpha^2).
\end{aligned}$$

Thus, the vortex stretching due to the beta effect enters first, at a lower order than wave effects in this scaling, if $\alpha^2 = \mathrm{o}(\beta/\epsilon)$. Since our main concern here is WMFI dynamics, we ignore the beta effect in what follows, although there apparently are flow regimes in which the beta effect either competes with WMFI dynamics, or dominates it, depending on the relative magnitudes of β/ϵ and α^2.

4.4. Noether symmetries, momentum balance and potential vorticity conservation

Noether's theorem associates conservation laws to continuous symmetries of HP. See, e.g., [39] for a clear discussion. Invariance of $\bar{\mathcal{L}}$ in Eq. (3.11) under time translations leads to the energy conservation relation

$$\frac{\partial W}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot \boldsymbol{S} = \mathrm{O}(\alpha^2\epsilon), \tag{4.17}$$

where W is the total energy density, given by

$$W/D = \tfrac{1}{2} |\bar{\boldsymbol{u}}_{\mathrm{L}}|^2 + \alpha^2 |\boldsymbol{a}|^2 \tilde{\omega}^2 + \rho g z + \mathrm{O}(\alpha^2\epsilon) \tag{4.18}$$

and S is the energy flux

$$S^j = \bar{u}^j_{\mathrm{L}}(W + D\overline{p}_{0\mathrm{L}}) - \alpha^2 \mathrm{i} D\omega(a^j b^* - a^{*j} b) + \alpha^2\epsilon \frac{\partial p_0}{\partial x^l}\left[a^{*l}\frac{\partial a^j}{\partial \epsilon t} + a^l \frac{\partial a^{*j}}{\partial \epsilon t}\right] + \mathrm{O}(\alpha^2\epsilon). \tag{4.19}$$

The quantity $\overline{p}_{0\mathrm{L}}$ is the Lagrangian mean of the pressure component p_0.

$$\overline{p}_{0\mathrm{L}} = \overline{p_0(\boldsymbol{x}+\alpha\boldsymbol{\xi}, t)} = p_0 + \alpha^2 \frac{1}{2}\frac{\partial^2 p_0}{\partial x^i \partial x^j}(a^i a^{*j} + a^{*i}a^j) + \mathrm{O}(\alpha^4\epsilon^2). \tag{4.20}$$

With an error of $\mathrm{O}(\alpha^2\epsilon)$, $\overline{p}_{0\mathrm{L}}$ is also the Lagrangian mean of the total pressure given in Eq. (2.11)

If $\overline{\mathcal{L}}$ in (3.11) did not contain the terms $\bar{\boldsymbol{u}}_{\mathrm{L}} \cdot (\boldsymbol{\Omega} \times \boldsymbol{x})$ and $\rho g z$ it would be invariant under spatial translations, which would lead to momentum conservation. Instead, we have the momentum *balance* relation

$$\frac{\partial G_i}{\partial \epsilon t} + \frac{\partial T^j_i}{\partial \epsilon x^j} + \frac{D}{\epsilon}\rho g \delta_{i3} + \frac{D}{\epsilon}(\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_{\mathrm{L}})_i = \mathrm{O}(\alpha^2\epsilon). \tag{4.21}$$

Here G_i is the total momentum density,

$$G_i = m_i + \alpha^2 N k_i = D[(\bar{u}_{\mathrm{L}})_i + (\boldsymbol{\Omega} \times \boldsymbol{x})_i], \tag{4.22}$$

consisting of the sum of the Eulerian momentum density in (3.16) and the wave momentum density $\alpha^2 N\boldsymbol{k}$, and T^j_i is the momentum stress tensor, given by

$$T^j_i = G_i \bar{u}^j_{\mathrm{L}} + D\overline{p}_{0\mathrm{L}}\delta^j_i - \alpha^2 \mathrm{i} D k_i(a^j b^* - a^{*j} b) - \alpha^2\epsilon \frac{\partial p_0}{\partial x^l}\left[a^{*l}\frac{\partial a^j}{\partial \epsilon x^i} + a^l \frac{\partial a^{*j}}{\partial \epsilon x^i}\right] + \mathrm{O}(\alpha^2\epsilon). \tag{4.23}$$

Since $\bar{\boldsymbol{u}}_{\mathrm{L}}$ has zero divergence at this order, the divergence of Eq. (4.21) gives an equation for the pressure which has no partial time derivatives. The operator acting on $\overline{p}_{0\mathrm{L}}$ is a small deformation of an elliptic operator, as we exhibit explicitly in the next section, and is thus invertible. In this sense, the theory is "balanced". Also, if the matrix of pressure derivatives $\partial p_0/(\partial x^j \partial x^l)$ initially depends only on $\epsilon\boldsymbol{x}$, ϵt, then it will remain so, provided the initial buoyancy satisfies (2.24) and the vertical component of $\bar{\boldsymbol{u}}_{\mathrm{L}}$ satisfies (2.26) throughout the time evolution. The terms proportional to α^2 in Eq. (4.23) for the total stress tensor are the radiation stress terms, see, e.g., [15]

The action $\overline{\mathcal{L}}$ is also invariant under a shift of Lagrangian labels along level surfaces of ρ effected by the vector field $\epsilon^{ABC}\mu_C(\partial\rho/\partial l^B)(\partial/\partial l^A)$, where μ_C is a constant, but otherwise arbitrary vector. The associated local conservation relation is

$$\frac{\partial}{\partial \epsilon t}\left[\epsilon^{ABC}\frac{\partial \rho}{\partial l^B}(D^{-1})^k_A m_k\right] + \frac{\partial}{\partial \epsilon x^i}\left[\epsilon^{ABC}\frac{\partial \rho}{\partial l^B}\left(\bar{u}^i_{\mathrm{L}}(D^{-1})^k_A m_k - D(D^{-1})^i_A \frac{\delta \overline{\mathcal{L}}}{\delta D}\right)\right] = \mathrm{O}(\alpha^2\epsilon). \tag{4.24}$$

Further manipulations of (4.24) yield the conservation law

$$\frac{\partial Dq}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot Dq\bar{\boldsymbol{u}}_{\mathrm{L}} = \mathrm{O}(\alpha^2\epsilon), \tag{4.25}$$

where q is the potential vorticity given by

$$q = \frac{1}{D}\frac{\partial \rho}{\partial \boldsymbol{x}} \cdot \left(\frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \frac{\boldsymbol{m}}{D}\right). \tag{4.26}$$

The continuity equation for D (3.7) then implies the convection relation for the potential vorticity q, namely $\mathrm{d}q/\mathrm{d}\epsilon t = \mathrm{O}(\alpha^2\epsilon)$. Eq. (4.25), or the closely related convection relation for q, can also be deduced from Kelvin's theorem, as we show in Section 4.6.

4.5. Solving for p_0

Differentiating the momentum balance equation (4.21) with respect to ϵx^i and summing to form the divergence gives, after setting $D = 1$, the following fourth-order equation for p_0, the Eulerian mean pressure:

$$\begin{aligned}&\left\{\frac{\partial^2}{\partial(\epsilon x^i)^2} + \alpha^2\epsilon^2\left[a^j a^{*l}\frac{\partial^4}{\partial(\epsilon x^i)^2\partial\epsilon x^j\partial\epsilon x^l} + \left(\frac{\partial}{\partial\epsilon x^i}(a^j a^{*l})\right)\frac{\partial^3}{\partial\epsilon x^i\partial\epsilon x^j\partial\epsilon x^l}\right.\right.\\&\left.\left.-\left(\frac{\partial}{\partial\epsilon x^j}\left(a^{*l}\frac{\partial a^j}{\partial\epsilon x^i} + a^l\frac{\partial a^{*j}}{\partial\epsilon x^i}\right)\right)\frac{\partial^2}{\partial\epsilon x^i\partial\epsilon x^l}\right]\right\}p_0\\&= -\frac{1}{\epsilon}g\frac{\partial\rho}{\partial\epsilon z} + \frac{2}{\epsilon}\boldsymbol{\Omega}\cdot(\frac{\partial}{\partial\epsilon\boldsymbol{x}}\times\bar{\boldsymbol{u}}_{\mathrm{L}}) - \frac{\partial\bar{u}^i_{\mathrm{L}}}{\partial\epsilon x^j}\frac{\partial\bar{u}^j_{\mathrm{L}}}{\partial\epsilon x^i}\\&+\alpha^2 \mathrm{i}\frac{\partial^2}{\partial\epsilon x^i\partial\epsilon x^j}\left[k^i(a^j b^* - a^{*j}b)\right] + \mathrm{O}(\alpha^2\epsilon).\end{aligned}\tag{4.27}$$

For times up to order $\mathrm{O}(1/\epsilon)$, Eq. (2.26) for the vertical Lagrangian mean velocity and Eq. (2.4) for the evolution of the mean displacement imply – after using the Mean Value Theorem on Eq. (2.4) and inverting the (nonsingular) Jacobian matrix $(D^{-1})^i_A$ – the following expressions for the Lagrangian fluid labels in terms of the Eulerian variables:

$$l^3 = z + \chi^3(\epsilon\boldsymbol{x}, \epsilon t) \quad\text{or}\quad \epsilon l^3 = \epsilon z + \epsilon\chi^3(\epsilon\boldsymbol{x}, \epsilon t) \qquad l^a = (1/\epsilon)\chi^a(\epsilon\boldsymbol{x}, \epsilon t), \quad a = 1, 2 \tag{4.28}$$

for certain functions χ^A, $A = 1, 2, 3$, of order O(1). Consequently, Eq. (2.24) for ρ and (2.25) for p_0 can be written in Eulerian form as

$$\rho = (1/\epsilon)(\epsilon z)r'_{-1}(\epsilon z, \epsilon t) + r_0(\epsilon\boldsymbol{x}, \epsilon t) + \epsilon r_1(\epsilon\boldsymbol{x}, \epsilon t) + \cdots \tag{4.29}$$

and

$$p_0 = (1/\epsilon^2)(\epsilon z)^2\pi'_{-2}(\epsilon z, \epsilon t) + (1/\epsilon)\pi_{-1}(\epsilon\boldsymbol{x}, \epsilon t) + \pi_0(\epsilon\boldsymbol{x}, \epsilon t) + \epsilon\pi_1(\epsilon\boldsymbol{x}, \epsilon t) + \cdots \tag{4.30}$$

We expand the Lagrangian mean velocity $\bar{\boldsymbol{u}}_{\mathrm{L}}$ in powers of ϵ as

$$\bar{\boldsymbol{u}}_{\mathrm{L}} = \bar{\boldsymbol{u}}^{(0)}_{\mathrm{L}} + \epsilon\bar{\boldsymbol{u}}^{(1)}_{\mathrm{L}} + \epsilon^2\bar{\boldsymbol{u}}^{(2)}_{\mathrm{L}} + \cdots, \tag{4.31}$$

and solve Eq. (4.27) perturbatively for p_0 by equating coefficients of like powers of ϵ. There is no need to expand any quantities in powers of α^2. Moreover, because of the error already assumed in the equation for momentum balance, and hence in Eq. (4.27), there is no need to expand wave quantities in powers of ϵ either.

At each order in ϵ^n, $n = -2, -1, 0, 1, \ldots$, the solution for pressure requires only the inversion of the Laplacian. The lowest order is

$$\frac{\partial^2((\epsilon z)^2\pi'_{-2})}{\partial(\epsilon z)^2} = -g\frac{\partial((\epsilon z)r'_{-1})}{\partial\epsilon z}, \tag{4.32}$$

which upon the imposition of the boundary condition

$$\left.\frac{\partial((\epsilon z)^2\pi'_{-2})}{\partial\epsilon z}\right|_{z=0} = 0$$

leads to the hydrostatic balance relation

$$\frac{\partial((\epsilon z)^2\pi'_{-2})}{\partial\epsilon z} = -g\epsilon z r'_{-1}. \tag{4.33}$$

Eq. (4.33) can also be obtained, of course, from the motion equation (4.1). If the function ρ' of Eq. (2.24), and consequently the function p_0' of Eq. (2.25) had depended on ϵl^a, $a = 1, 2$, then π'_{-2} would have also depended on ϵx and ϵy. Upon enforcing the geostrophic balance at leading order we would have been led to the equations

$$\frac{\partial \pi'_{-2}}{\partial \epsilon x^i} = 0, \quad i = 1, 2. \tag{4.34}$$

This demonstrates that π'_{-2}, and hence r'_{-1}, cf. Eq. (4.33) cannot depend on the transverse coordinates.

At order ϵ^{-1}, Eq. (4.27) for p_0 yields

$$\frac{\partial^2 \pi_{-1}}{\partial (\epsilon x^i)^2} = -g\frac{\partial r_0}{\partial \epsilon z} + 2\boldsymbol{\Omega} \cdot \left(\frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \bar{\boldsymbol{u}}_{\mathrm{L}}^{(0)}\right). \tag{4.35}$$

This equation for π_{-1}, together with the usual Neumann boundary conditions, is equivalent to the equations that result from the motion equation (4.1) at order $\mathrm{O}(1/\epsilon)$,

$$\frac{\partial \pi_{-1}}{\partial \epsilon z} = -g r_0,$$
$$\frac{\partial \pi_{-1}}{\partial \epsilon \boldsymbol{x}_{\mathrm{H}}} = -2\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_{\mathrm{L}}^{(0)}. \tag{4.36}$$

The quantity $\boldsymbol{x}_{\mathrm{H}}$ is the transverse component of the vector $\boldsymbol{x}$, as defined after Eq. (4.7).

Note that because the pressure equation (4.27) assumes an error of $\mathrm{O}(\alpha^2\epsilon)$, the fourth-order differential operator in square brackets multiplied by $\alpha^2\epsilon^2$ in Eq. (4.27) and acting on p_0 finally enters only at order ϵ^0, when it acts on $(\epsilon z)^2\pi'_{-2}$. At this order, we have a Poisson equation for π_0,

$$\begin{aligned}\frac{\partial^2 \pi_0}{\partial (\epsilon x^i)^2} = &-\alpha^2\Bigg[|a^3|^2\frac{\partial^4}{\partial(\epsilon z)^4} + \left(\frac{\partial |a^3|^2}{\partial \epsilon z}\right)\frac{\partial^3}{\partial(\epsilon z)^3} \\ &- \left(\frac{\partial}{\partial \epsilon x^j}\left(a^{*3}\frac{\partial a^j}{\partial \epsilon z} + a^3\frac{\partial a^{*j}}{\partial \epsilon z}\right)\right)\frac{\partial^2}{\partial(\epsilon z)^2}\Bigg](\epsilon z)^2\pi'_{-2} \\ &- g\frac{\partial r_1}{\partial \epsilon z} + 2\boldsymbol{\Omega}\cdot\left(\frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \bar{\boldsymbol{u}}_{\mathrm{L}}^{(1)}\right) - \frac{\partial \bar{u}_{\mathrm{L}}^{(0)i}}{\partial \epsilon x^j}\frac{\partial \bar{u}_{\mathrm{L}}^{(0)j}}{\partial \epsilon x^i} + \mathrm{i}\alpha^2\frac{\partial^2}{\partial \epsilon x^i \partial \epsilon x^j}[k^i(a^j b^* - a^{*j} b)],\end{aligned} \tag{4.37}$$

whose boundary conditions are obtained by evaluating the momentum balance relation (4.21) on the boundary. Namely,

$$\begin{aligned}\hat{\boldsymbol{n}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}}\pi_0 = &-n_3 g r_1 - 2\hat{\boldsymbol{n}} \cdot (\boldsymbol{\Omega} \times \bar{\boldsymbol{u}}_{\mathrm{L}}^{(1)}) - n_i \bar{u}_{\mathrm{L}}^{(0)j}\frac{\partial}{\partial \epsilon x^j}\bar{u}_{\mathrm{L}}^{(0)i} \\ &+ \alpha^2 \mathrm{i} n_i \frac{\partial}{\partial \epsilon x^j}[k^i(a^j b^* - a^{*j} b)] - \alpha^2 |a^3|^2 \hat{\boldsymbol{n}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}}\left(\frac{\partial^2 (\epsilon z)^2 \pi'_{-2}}{\partial(\epsilon z)^2}\right),\end{aligned} \tag{4.38}$$

where $\hat{\boldsymbol{n}}$ is the unit normal vector to the boundary, and both sides of the equation are evaluated at the boundary.

Remark. Since the derivative of ρ with respect to l^3 (or z) is assumed to be negative, the dispersion relation (3.24) implies that $\tilde{\omega}^2$ is positive definite. (The quantity $\sqrt{-g\partial\rho/\partial z}$ is Brunt–Väisällä frequency for stable oscillations under the restoring force of gravity.) In fact, as mentioned earlier, the Doppler-shifted wave frequency is assumed to be of order $\mathrm{O}(1)$. In combination with the dispersion relation (3.24) this implies that the projection of the curvature of an isobar onto a plane orthogonal to the wave vector is bounded, both above and below. This, of course, is consistent with the assumed forms of ρ and p_0, Eqs. (2.24) and (2.25), and the form of w, Eq. (2.26). However, it also affords a remarkable simplification for solvability of the WMFI system that the curvature of the pressure contours is bounded for times of order $O(1/\epsilon)$.

4.6. Kelvin's theorem and potential vorticity

Kelvin's theorem for this theory is found by integrating the motion equation (3.15) around a closed contour $\overline{\gamma}(\epsilon t)$ which moves with the Lagrangian mean velocity $\overline{\boldsymbol{u}}_{\mathrm{L}}$. (In fact, the motion equation is naturally in the form of Kelvin's theorem, when expressed in terms of variations of $\overline{\mathcal{L}}$ with respect to D and $\overline{\boldsymbol{u}}_{\mathrm{L}}$). Suppressing the error estimates, we calculate, as follows :

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\oint_{\overline{\gamma}(\epsilon t)}\frac{1}{D}\boldsymbol{m}\cdot\mathrm{d}\boldsymbol{x}&=\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\oint_{\overline{\gamma}(\epsilon t)}\left(\overline{\boldsymbol{u}}_{\mathrm{L}}-\alpha^2\frac{N}{D}\boldsymbol{k}+\boldsymbol{\Omega}\times\boldsymbol{x}\right)\cdot\mathrm{d}\boldsymbol{x}\\&=\oint_{\overline{\gamma}(\epsilon t)}\left[\left(\frac{\partial}{\partial\epsilon t}+\overline{\boldsymbol{u}}_{\mathrm{L}}\cdot\frac{\partial}{\partial\epsilon\boldsymbol{x}}\right)\frac{\boldsymbol{m}}{D}+\frac{m_j}{D}\frac{\partial\overline{u}_{\mathrm{L}}^{\,j}}{\partial\epsilon\boldsymbol{x}}\right]\cdot\mathrm{d}\boldsymbol{x}\\&=-\frac{g}{\epsilon}\oint_{\overline{\gamma}(\epsilon t)}\rho\hat{\boldsymbol{z}}\cdot\mathrm{d}\boldsymbol{x}=-\frac{g}{\epsilon}\int\int_{S(\overline{\gamma})}\boldsymbol{\nabla}\rho\times\hat{\boldsymbol{z}}\cdot\mathrm{d}\boldsymbol{S}.\end{aligned}\tag{4.39}$$

where $S(\overline{\gamma})$ is a surface with boundary $\overline{\gamma}$. Thus, a nonvertical gradient of buoyancy may create a *combination* of fluid circulation and weak wave activity. We may rewrite this equation as

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\oint_{\overline{\gamma}(\epsilon t)}(\overline{\boldsymbol{u}}_{\mathrm{L}}+\boldsymbol{\Omega}\times\boldsymbol{x})\cdot\mathrm{d}\boldsymbol{x}+\frac{g}{\epsilon}\oint_{\overline{\gamma}(\epsilon t)}\rho\hat{\boldsymbol{z}}\cdot\mathrm{d}\boldsymbol{x}=\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\oint_{\overline{\gamma}(\epsilon t)}\alpha^2\frac{N}{D}\boldsymbol{k}\cdot\mathrm{d}\boldsymbol{x}.\tag{4.40}$$

Consequently, if the wave component of the circulation is steady, it creates no Lagrangian mean-flow circulation. This is a version of the Charney–Drazin "nonacceleration" theorem discussed in the introduction, see, e.g., [6,7].

Applying the transport theorem (cf. [12]) to the right-hand side of Eq. (4.40) gives

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\oint_{\overline{\gamma}(\epsilon t)}(\overline{\boldsymbol{u}}_{\mathrm{L}}+\boldsymbol{\Omega}\times\boldsymbol{x})\cdot\mathrm{d}\boldsymbol{x}+\frac{g}{\epsilon}\oint_{\overline{\gamma}(\epsilon t)}\rho\hat{\boldsymbol{z}}\cdot\mathrm{d}\boldsymbol{x}\\&=\alpha^2\oint_{\overline{\gamma}(\epsilon t)}\left[\frac{\partial}{\partial\epsilon t}\frac{N}{D}\boldsymbol{k}-\overline{\boldsymbol{u}}_{\mathrm{L}}\times\left(\frac{\partial}{\partial\epsilon\boldsymbol{x}}\times\frac{N}{D}\boldsymbol{k}\right)\right]\cdot\mathrm{d}\boldsymbol{x}.\end{aligned}\tag{4.41}$$

Therefore, we may say more precisely that (upon evaluating $D=1$) a steady wave whose wave vector $\boldsymbol{k}$ is aligned with its wave action density gradient $\partial N/\partial(\epsilon\boldsymbol{x})$ creates no circulation of Lagrangian mean velocity.

Applying Stokes' theorem to Eq. (4.39) reconfirms that the potential vorticity q defined by Eq. (4.26) is convected with the Lagrangian mean velocity $\overline{\boldsymbol{u}}_{\mathrm{L}}$, i.e.,

$$\frac{\partial q}{\partial\epsilon t}+\overline{\boldsymbol{u}}_{\mathrm{L}}\cdot\frac{\partial q}{\partial\epsilon\boldsymbol{x}}\equiv\frac{\mathrm{d}q}{\mathrm{d}\epsilon t}=0\tag{4.42}$$

Consequently, we have the conserved quantities ("Casimirs"),

$$C_\Phi=\int\mathrm{d}^3x\,D\Phi(q,\rho)=\mathrm{const.}\tag{4.43}$$

for any fuction Φ.

Remark. The leading order WMFI circulation theorem (4.41) extends to arbitrary fluid contours the circulation theorem of Grimshaw [23] for horizontal fluid contours – in the incompressible limit with the Boussinesq approximation, and when viscosity is absent. In the same limit, Eq. (4.42) extends the equation for potential vorticity dynamics in [23] to a conservation law for potential vorticity on mean fluid parcels, thereby leading to the Casimirs in Eq. (4.43).

4.7. Summary of WMFI equations to order $O(\alpha^2)$

Suppressing the asymptotic order estimates, we may summarize the self-consistent WMFI equations of this section, as follows (with $D = 1$):

$$\left(\frac{\partial}{\partial \epsilon t} + \bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \frac{\partial}{\partial \epsilon \boldsymbol{x}}\right)\bar{\boldsymbol{u}}_{\mathrm{L}} - \frac{1}{\epsilon}\left(\bar{\boldsymbol{u}}_{\mathrm{L}} \times 2\boldsymbol{\Omega} - \rho g \hat{\boldsymbol{z}} - \frac{\partial p_0}{\partial \boldsymbol{x}}\right)$$
$$= -\alpha^2\left[N\frac{\partial \tilde{\omega}}{\partial \epsilon \boldsymbol{x}} + \mathrm{i}\boldsymbol{k}\left(\frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot (\boldsymbol{a}^* b - \boldsymbol{a} b^*)\right)\right]$$

with

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = 0, \quad \frac{\partial N}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon x} \cdot \left(\bar{\boldsymbol{u}}_{\mathrm{L}} N + \mathrm{i}(\boldsymbol{a}^* b - \boldsymbol{a} b^*)\right) = 0, \quad \frac{\partial \boldsymbol{k}}{\partial \epsilon t} + \frac{\partial \omega}{\partial \epsilon \boldsymbol{x}} = 0. \tag{4.44}$$

The quantity N (wave action density) is defined by

$$N = 2|\boldsymbol{a}|^2\tilde{\omega} + 2\mathrm{i}\boldsymbol{\Omega} \cdot (\boldsymbol{a} \times \boldsymbol{a}^*). \tag{4.45}$$

At a given time step, the diagnostic variables, p_0, b, $\boldsymbol{a}$ and $\tilde{\omega} = \omega - \bar{\boldsymbol{u}}_{\mathrm{L}} \cdot \boldsymbol{k}$, are determined from the current value of N and the constraints (again with $D = 1$),

$$\frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot \bar{\boldsymbol{u}}_{\mathrm{L}} = 0, \qquad \boldsymbol{k} \cdot \boldsymbol{a} = 0, \qquad \tilde{\omega}^2 \boldsymbol{a} + 2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega} \times \boldsymbol{a}) - \mathrm{i}b\boldsymbol{k} - \left(\boldsymbol{a} \cdot \frac{\partial}{\partial \boldsymbol{x}}\right)\frac{\partial p_0}{\partial \boldsymbol{x}} = 0 \tag{4.46}$$

up to an overall phase in both $\boldsymbol{a}$ and b.

5. Hamiltonian structure

We now pass to the Hamiltonian formulation of the leading order WMFI equations (4.44). For this, the relation

$$\pi_A \frac{\partial l^A}{\partial t} - \alpha^2 N \frac{\partial \phi}{\partial \epsilon t} = \boldsymbol{m} \cdot \bar{\boldsymbol{u}}_{\mathrm{L}} + \alpha^2 N \omega \tag{5.1}$$

is useful in Legendre-transforming $\overline{\mathcal{L}}$ in (3.11) to find the WMFI Hamiltonian, $\overline{\mathcal{H}}$. Then we transform the Poisson bracket in the mean-flow canonically conjugate variables π_A and $l^A(\boldsymbol{x}, t)$, $A = 1, 2, 3$, to the noncanonical Eulerian fluid variables, $D = \det \nabla l^A$, $\rho(l^A)$ and $\boldsymbol{m}$ given in Eq. (3.16), by using the chain rule for functional derivatives. We also transform the Poisson bracket in the wave canonically conjugate variables $-\alpha^2 N$ and ϕ/ϵ to the noncanonical wave variables N and $\boldsymbol{p}$, where $\boldsymbol{p}$ is the wave pseudomomentum density, $\boldsymbol{p} \equiv N\boldsymbol{k}$, again by using the chain rule for functional derivatives. This yields the ideal WMFI equations in Lie–Poisson Hamiltonian form [1,31], in terms of the Eulerian fluid and wave variables. The ideal wave mean-flow system (4.44) turns out to be a Lie–Poisson Hamiltonian fluid system. Thus, its (relative) equilibrium solutions are critical points of a constrained energy and the energy-Casimir method may be applied to determine Lyapunov stability conditions for these equilibria, as done in, e.g., [32].

Passing from the constrained Lagrangian (3.11) for the WMFI equations via the Legendre transformation – facilitated by relation (5.1) – yields the following constrained WMFI Hamiltonian:

$$\begin{aligned}\overline{\mathcal{H}} = \int \mathrm{d}^3x \Bigg\{ \Bigg[& \frac{1}{2D}|\boldsymbol{m} + \alpha^2\boldsymbol{p} - D(\boldsymbol{\Omega}\times\boldsymbol{x}|^2 + D\rho g z \\ & + \frac{\alpha^2 D}{4|\boldsymbol{a}|^2}\left(\frac{N}{D} - 2\mathrm{i}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*)\right)\Bigg] + \frac{\alpha^2 i D}{N}(b\boldsymbol{p}\cdot\boldsymbol{a}^* - b^*\boldsymbol{p}\cdot\boldsymbol{a}) \\ & + (D-1)\left[p_0 + \frac{\alpha^2}{2}\frac{\partial^2 p_0}{\partial x^i \partial x^j}(a^i a^{*j} + a^j a^{*i})\right] \\ & - \alpha^2\epsilon\left(a^m \frac{\partial p_0}{\partial x^m}\frac{\partial a^{*j}}{\partial \epsilon x^j} + a^{*m}\frac{\partial p_0}{\partial x^m}\frac{\partial a^j}{\partial \epsilon x^j}\right) + \mathrm{O}(\alpha^2\epsilon, \alpha^4)\Bigg\}. \end{aligned} \tag{5.2}$$

Actually, $\overline{\mathcal{H}}$ is a Routhian [42] – the diagnostic variables $\boldsymbol{a}$, b, and p_0 (as opposed to the prognostic variables l^A and ϕ) are not Legendre-transformed, since they have no canonically conjugate momenta at this order. Thus, they remain as Lagrange multipliers which impose the *same* constraints as they did before, in Eqs. (4.46). Perhaps not unexpectedly, evaluating $\overline{\mathcal{H}}$ on its constraint manifold by setting $D = 1$, using (4.46) and the definitions (3.16) and (3.17), recovers the conserved WMF energy in (4.18).

The canonical Poisson bracket in terms of (π_A, l^A) and $(-\alpha^2 N, \phi/\epsilon)$, which follows from HP with Lagrangian (3.10) is given by

$$\begin{aligned} &\{F, G\}(\pi_A, l^A, -\alpha^2 N, \phi/\epsilon) \\ &= -\int \mathrm{d}^3x \left[\frac{\delta F}{\delta \pi_A}\frac{\delta G}{\delta l^A} - \frac{\delta G}{\delta \pi_A}\frac{\delta F}{\delta l^A} - \frac{\delta F}{\delta(\alpha^2 N)}\frac{\delta G}{\delta \phi/\epsilon} + \frac{\delta F}{\delta \phi/\epsilon}\frac{\delta G}{\delta(\alpha^2 N)}\right]. \end{aligned} \tag{5.3}$$

The WMFI equations (4.44) now follow in Hamilton's canonical form upon using $\overline{\mathcal{H}}$ and the canonically conjugate pairs (π_A, l^A) and $(-\alpha^2 N, \phi/\epsilon)$. However, since $\overline{\mathcal{H}}$ is expressed in the Eulerian fluid variables $\boldsymbol{m}$, D, ρ, and the wave variables $\boldsymbol{p}$, N, we transform the canonical Poisson bracket (5.3) into the variables that appear in the Hamiltonian, in order to simplify the subsequent manipulations. (These variables Poisson-commute amongst themselves under the canonical Poisson bracket, so this transformation is a Poisson map.)

Specifically, the definitions $D = \det \boldsymbol{\nabla} l^A$ and $\boldsymbol{m} = -\pi_A \boldsymbol{\nabla} l^A$ in (3.6) and (3.18) and $\boldsymbol{p} = N\boldsymbol{k}$ allow one to use the chain rule to transform the canonical Poisson bracket into a sum, consisting of the Lie–Poisson bracket in terms of variables $\boldsymbol{m}$, D and ρ that is discussed, e.g., in [1,31], and another similar Lie–Poisson bracket in terms of $\boldsymbol{p}$ and N. Namely,

$$\begin{aligned} \{F, G\} = -\Omega\epsilon \int \mathrm{d}^3x \Bigg[& \frac{\delta F}{\delta m_i}\left((\partial_j m_i + m_j \partial_i)\frac{\delta G}{\delta m_j} + D\partial_i \frac{\delta G}{\delta D} - \frac{\delta G}{\delta \rho}\partial_i \rho\right) \\ & + \frac{\delta F}{\delta D}\partial_j\left(D\frac{\delta G}{\delta m_j}\right) + \frac{\delta F}{\delta \rho}\frac{\delta G}{\delta m_j}\partial_j \rho\Bigg] \\ - \frac{\epsilon}{\alpha^2}\int \mathrm{d}^3x & \left[\frac{\delta F}{\delta p_i}\left((\partial_j p_i + p_j \partial_i)\frac{\delta G}{\delta p_j} + N\partial_i \frac{\delta G}{\delta N}\right) + \frac{\delta F}{\delta N}\partial_j\left(N\frac{\delta G}{\delta p_j}\right)\right], \end{aligned} \tag{5.4}$$

where we have integrated by parts and the partial derivative $\partial_j = \partial/\partial\epsilon x^j$, $j = 1, 2, 3$, operates on all terms it multiplies to its right. This Lie–Poisson bracket satisfies the Jacobi identity

$$\{E, \{F, G\}\} + \{F, \{G, E\}\} + \{G, \{E, F\}\} = 0 \tag{5.5}$$

for any functionals E, F and G of $(\boldsymbol{m}, D, \rho, \boldsymbol{p}, N)$, simply because (5.5) is a variable transform of the Jacobi identity for the canonical Poisson bracket (5.3). The first Lie–Poisson bracket in Eq. (5.4) is defined on the dual of the semidirect-product Lie algebra of vector fields and the direct sum of functions and densities. Dual coordinates are: $\boldsymbol{m}$, dual to vector fields; D, dual to functions; and ρ, dual to densities. See, e.g., [29,30]. The other Lie–Poisson bracket in the sum is similar. Its dual coordinates are: $\boldsymbol{p}$, dual to vector fields; N, dual to functions. Each Lie–Poisson bracket in the sum satisfies the Jacobi identity separately. This Lie–Poisson structure is identical in form to that found by Holm and Kupershmidt [28] for the two-fluid theory of superfluid He^4 in the absence of vortices. The difference is in interpretation. In the two-fluid theory of superfluids, there are two interpenetrating fluids. The normal fluid carries heat and mass, while the superfluid carries only mass. In the two-fluid interpretation of the WMFI theory, the mean flow carries volume and mass, while the "wave fluid" carries only wave action. In both cases there is an S^1 order parameter, a phase. A similar analogy – between superfluid models and an ensemble of high frequency sound waves in a compressible fluid – is discussed by Putterman and Roberts [41].

The variational derivatives of the constrained Hamiltonian (5.2) may be determined from the coefficients in the following expression :

$$
\begin{aligned}
\delta\overline{\mathcal{H}} = \int \mathrm{d}^3x \Bigg\{ & \left[-\frac{1}{2}|\overline{\boldsymbol{u}}_{\mathrm{L}}|^2 - \overline{\boldsymbol{u}}_{\mathrm{L}} \cdot (\boldsymbol{\Omega}\times\boldsymbol{x}) + \rho g z + \overline{p} \right] \delta D + D g z \delta\rho \\
& + \overline{\boldsymbol{u}}_{\mathrm{L}} \cdot \delta\boldsymbol{m} - (1-D)\delta p_0 + \alpha^2 \left[\tilde{\omega} - \frac{\mathrm{i}D}{N}(b\boldsymbol{k}\cdot\boldsymbol{a}^* - b^*\boldsymbol{k}\cdot\boldsymbol{a}) \right] \delta N \\
& + \alpha^2 \left[\overline{\boldsymbol{u}}_{\mathrm{L}} + \frac{\mathrm{i}D}{N}(b\boldsymbol{a}^* - b^*\boldsymbol{a}) \right] \cdot \delta\boldsymbol{p} + \mathrm{i}\alpha^2 D(\delta b\,\boldsymbol{k}\cdot\boldsymbol{a}^* - \delta b^*\boldsymbol{k}\cdot\boldsymbol{a}) \\
& - \alpha^2 \left[\delta\boldsymbol{a}^* \cdot \left(D\tilde{\omega}^2\boldsymbol{a} + 2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega}\times\boldsymbol{a}) - \mathrm{i}Db\boldsymbol{k} - D\left(\boldsymbol{a}\cdot\frac{\partial}{\partial\boldsymbol{x}}\right)\frac{\partial p_0}{\partial\boldsymbol{x}} \right) + \text{c.c.} \right] + \mathrm{O}(\alpha^2\epsilon) \Bigg\},
\end{aligned} \tag{5.6}
$$

where the quantity $\overline{p}$ is given by

$$
\overline{p} = p_0 + \frac{\alpha^2}{2}\frac{\partial^2 p_0}{\partial x^i \partial x^j}(a^i a^{*j} + a^j a^{*i}) - \alpha^2(\tilde{\omega}^2|\boldsymbol{a}|^2 + 2\mathrm{i}\tilde{\omega}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*)) + \mathrm{O}(\alpha^2\epsilon). \tag{5.7}
$$

The corresponding equations of motion in Hamiltonian form, correct through order $\mathrm{O}(\alpha^2)$, are given by

$$
\begin{aligned}
\frac{\partial m_i}{\partial \epsilon t} &= \frac{1}{\epsilon}\{m_i, \overline{\mathcal{H}}\} = -(\partial_j m_i + m_j \partial_i)\frac{\delta\overline{\mathcal{H}}}{\delta m_j} + \frac{\delta\overline{\mathcal{H}}}{\delta\rho}\partial_i\rho - D\partial_i\frac{\delta\overline{\mathcal{H}}}{\delta D} \\
&= -(\partial_j m_i + m_j\partial_i)\overline{u}^j_{\mathrm{L}} + Dgz\partial_i\rho - D\partial_i\left(-\frac{1}{2}|\overline{\boldsymbol{u}}_{\mathrm{L}}|^2 - \overline{\boldsymbol{u}}_{\mathrm{L}}\cdot(\boldsymbol{\Omega}\times\boldsymbol{x}) + \rho g z + \overline{p}\right), \\
\frac{\partial D}{\partial \epsilon t} &= \frac{1}{\epsilon}\{D, \overline{\mathcal{H}}\} = -\partial_j D\frac{\delta\overline{\mathcal{H}}}{\delta m_j} = -\partial_j D\overline{u}^j_{\mathrm{L}}, \\
\frac{\partial \rho}{\partial \epsilon t} &= \frac{1}{\epsilon}\{\rho, \overline{\mathcal{H}}\} = -\frac{\delta\overline{\mathcal{H}}}{\delta m_j}\partial_j\rho = -\overline{u}^j_{\mathrm{L}}\partial_j\rho, \\
\frac{\partial N}{\partial \epsilon t} &= \frac{1}{\epsilon}\{N, \overline{\mathcal{H}}\} = -\frac{1}{\alpha^2}\partial_j\left(N\frac{\delta\overline{\mathcal{H}}}{\delta p_j}\right) = -\partial_j(N\overline{u}^j_{\mathrm{L}} + 2D\,\mathrm{Im}(b^*a^j)), \\
\frac{\partial p_j}{\partial \epsilon t} &= \frac{1}{\epsilon}\{p_j, \overline{\mathcal{H}}\} = -\frac{1}{\alpha^2}(\partial_j p_i + p_j\partial_i)\frac{\delta\overline{\mathcal{H}}}{\delta p_j} - \frac{1}{\alpha^2}N\partial_i\frac{\delta\overline{\mathcal{H}}}{\delta N} \\
&= -(\partial_j p_i + p_j\partial_i)\left(\overline{u}^j_L + \frac{2D}{N}\mathrm{Im}(b^*a^j)\right) - N\partial_i\left(\tilde{\omega} - \frac{2D}{N}\mathrm{Im}(b^*\boldsymbol{k}\cdot\boldsymbol{a})\right)
\end{aligned} \tag{5.8}
$$

with constraints determined by the other stationarity conditions,

$$\begin{aligned}
\delta p_0:&\quad D = 1 + \mathrm{O}(\alpha^2\epsilon^2),\\
\delta b^*:&\quad \boldsymbol{k}\cdot\boldsymbol{a} = \mathrm{O}(\epsilon),\\
\delta a^*:&\quad \tilde{\omega}^2\boldsymbol{a} + 2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega}\times\boldsymbol{a}) - ib\boldsymbol{k} - \left(\boldsymbol{a}\cdot\frac{\partial}{\partial\boldsymbol{x}}\right)\frac{\partial p_0}{\partial\boldsymbol{x}} = \mathrm{O}(\epsilon).
\end{aligned} \tag{5.9}$$

So $\overline{\mathcal{H}}$ is stationary for the diagnostic variables and generates the evolution of the prognostic variables. Straightforward manipulations simplify the last equation in (5.8) to $\partial\boldsymbol{k}/\partial\epsilon t = -\partial\omega/\partial\epsilon\boldsymbol{x}$, which is the wave transport equation.

Note that evaluating the pressure $\overline{p}$ in Eq. (5.7) on the constraint manifold (5.9) gives $\overline{p} = p_0 + \mathrm{O}(\alpha^2\epsilon)$. Then suppressing order estimates in Eqs. (5.8) and (5.9) yields the WMFI equations (4.44) and (4.46) in Lie–Poisson Hamiltonian form in terms of $\boldsymbol{m}$, ρ, N and $\boldsymbol{k}$.

5.1. Interpretation of the wave pseudomomentum density

The wave pseudomomentum density $\boldsymbol{p} = N\boldsymbol{k}$ is a subsystem momentum density in two important defining senses. First, under the Lie–Poisson bracket of Eq. (5.4), the total wave momentum with components $P_i = \int \alpha^2 N k_i\, \mathrm{d}^3x$ generates an Eulerian spatial shift of the wave properties, namely the phase ϕ/ϵ and its canonically conjugate momentum density $-\alpha^2 N$, i.e.,

$$\{P_i, \phi\} = \frac{\partial\phi}{\partial x^i}, \qquad \{P_i, \alpha^2 N\} = \frac{\partial\alpha^2 N}{\partial x^i}, \tag{5.10}$$

while leaving the fluid variables $\boldsymbol{m}$, D, ρ invariant. Likewise, the total Eulerian momentum components $M_i = \int m_i\, \mathrm{d}^3x$ generate the corresponding Eulerian spatial shifts of the fluid properties $(\boldsymbol{m}, D, \rho)$ while leaving the wave variables $(-\alpha^2 N, \phi/\epsilon)$ invariant. This can also be seen by using the definition $\boldsymbol{m} = -\pi_A \nabla l^A$ in the canonical Poisson bracket (5.3) to show that M_i generates the infinitesimal spatial shifts,

$$\{M_i, \pi_A\} = \frac{\partial\pi_A}{\partial x^i}, \qquad \{M_i, l^A\} = \frac{\partial l^A}{\partial x^i}, \tag{5.11}$$

while leaving the wave quantities $(-\alpha^2 N, \phi/\epsilon)$ invariant. Of course, then, the total momentum components, $\int (m_i + \alpha^2 N k_i)\, \mathrm{d}^3x$ generate Eulerian spatial shifts in the ith direction for the *entire* wave mean-flow system, i.e., in all these variables together.

Second, it is clear that

$$\alpha^2 N\boldsymbol{k}\cdot \mathrm{d}\boldsymbol{x} = \alpha^2 N\, \mathrm{d}\phi/\epsilon. \tag{5.12}$$

Thus, the quantity $\alpha^2 N\boldsymbol{k}\cdot\mathrm{d}\boldsymbol{x}$ is the canonical action one-form density in the phase space of wave properties which leads to the symplectic two-form

$$\overline{\Omega}_{\mathrm{wave}} = \alpha^2\, \mathrm{d}(Nk_i)\wedge \mathrm{d}x^i = \alpha^2\, \mathrm{d}N\wedge \mathrm{d}\phi/\epsilon. \tag{5.13}$$

Moreover, by Eq. (3.18) which relates the Eulerian momentum density $\boldsymbol{m}$ to the momentum $\pi_A \equiv \delta\overline{\mathcal{L}}/\delta l^A{}_{,t}$ canonically conjugate to the Lagrangian fluid label l^A, we have

$$D(\overline{\boldsymbol{u}}_{\mathrm{L}} + (\boldsymbol{\Omega}\times\boldsymbol{x}))\cdot \mathrm{d}\boldsymbol{x} = (\boldsymbol{m} + \alpha^2 N\boldsymbol{k})\cdot\mathrm{d}\boldsymbol{x} = -\pi_A\, \mathrm{d}l^A + \alpha^2 N\, \mathrm{d}\phi/\epsilon. \tag{5.14}$$

Consequently, the wave pseudomomentum density $\boldsymbol{p} = N\boldsymbol{k}$ and the Eulerian momentum density $\boldsymbol{m}$ enter the total canonical action one-form density (5.14) on precisely the same footing and describe complementary aspects of the wave-fluid system.

5.2. Relative equilibria and critical points of constrained energy

Under the Lie–Poisson bracket (5.4) for the WMFI dynamics, the infinitesimal transformation generated by the conserved quantities C_Φ in (4.43) leaves invariant the Eulerian fluid variables $\boldsymbol{m}$, D, ρ, N and $\boldsymbol{k}$, since C_Φ Poisson-commutes with all of these variables for an arbitrary function, Φ, of ρ and q. The corresponding infinitesimal canonical transformation (gauge transformation) of the Lagrangian fluid labels $l^A(\boldsymbol{x}, t)$, $A = 1, 2, 3$, is given by

$$\{C_\Phi, l^A\} = \tilde{\boldsymbol{v}} \cdot \boldsymbol{\nabla} l^A \quad \text{with } \tilde{\boldsymbol{v}} = D^{-1}\boldsymbol{\nabla}\rho \times \boldsymbol{\nabla}\frac{\partial \Phi}{\partial q} \quad \text{and } D = 1. \tag{5.15}$$

Thus, C_Φ generates a volume-preserving shift in the Lagrangian fluid labels along intersections of level surfaces of density ρ and potential vorticity q, that also leaves invariant the fluid's Eulerian momentum density, $\boldsymbol{m}$, and wave variables, N and $\boldsymbol{k}$. The corresponding relative equilibrium flow of the WMFI equations is given by $\bar{\boldsymbol{u}}_{\mathrm{Le}} = \tilde{\boldsymbol{v}}$ with $\tilde{\boldsymbol{v}}$ in (5.15).

The relative equilibrium solutions for WMFI dynamics are critical points of the sum $H_C = \overline{\mathcal{H}} + \nu \int \mathrm{d}^3x\, N + C_\Phi$, in which the function Φ is related to the Bernoulli function for the equilibrium solution and ν is a constant. Two obvious classes of equilibria are those for which there is either no wave amplitude, or the Lagrangian mean velocity is constant. In these cases, one studies the effect of either small amplitude waves on an equilibrium mean flow, or weak shear on an equilibrium wave field. The stability of these, and other relative equilibrium solutions may be investigated by using constrained energy methods similar to those developed for Euler's equations in the Boussinesq approximation in [2].

6. Higher-order corrections

6.1. Averaging in HP and the equations of motion

Here we extend the averaged HP to include terms multiplied by $\alpha^2\epsilon$ and $\alpha^2\epsilon^2$. The error in the action is then of order $\mathrm{O}(\alpha^4)$ which, as explained in the paragraph containing Eq. (3.10), gives an error of order $\mathrm{O}(\alpha^4\epsilon)$ in the equations of motion. At this order, the averaged expressions for the kinetic and potential energy are exact. The only truncation occurs in the incompressibility constraint imposed by the pressure as a Lagrange multiplier, in which we ignore terms of order $\mathrm{O}(\alpha^4)$. The averaged action $\bar{\mathcal{L}}$ is given by, cf. Eq. (3.10),

$$\begin{aligned}
\bar{\mathcal{L}} = \int \mathrm{d}t \int \mathrm{d}^3x \Bigg\{ D\Bigg[&\tfrac{1}{2}|\bar{\boldsymbol{u}}_{\mathrm{L}}|^2 + \alpha^2|\boldsymbol{a}|^2\tilde{\omega}^2 - \alpha^2\epsilon\mathrm{i}\tilde{\omega}\left(\boldsymbol{a}\cdot\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t} - \boldsymbol{a}^*\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right) \\
&+\alpha^2\epsilon^2\left|\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right|^2 - \rho\left(l^A(\boldsymbol{x}, t)\right) gz + \bar{\boldsymbol{u}}_{\mathrm{L}}\cdot(\boldsymbol{\Omega}\times\boldsymbol{x}) + 2\mathrm{i}\alpha^2\tilde{\omega}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*) \\
&+\alpha^2\epsilon\boldsymbol{\Omega}\cdot\left(\boldsymbol{a}\times\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t} + \boldsymbol{a}^*\times\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right)\Bigg] \\
&+(1-D)\left[p_0 + \frac{\alpha^2}{2}\frac{\partial^2 p_0}{\partial x^i\partial x^j}(a^i a^{*j} + a^j a^{*i}) + \alpha^2\left(\mathrm{i}\boldsymbol{k}b + \epsilon\frac{\partial b}{\partial\epsilon\boldsymbol{x}}\right)\cdot\boldsymbol{a}^*\right. \\
&\left.+\alpha^2\left(-\mathrm{i}\boldsymbol{k}b^* + \epsilon\frac{\partial b^*}{\partial\epsilon\boldsymbol{x}}\right)\cdot\boldsymbol{a}\right] + \alpha^2 p_0\left[\mathrm{i}\epsilon\boldsymbol{k}\cdot\frac{\partial}{\partial\epsilon\boldsymbol{x}}\times(\boldsymbol{a}\times\boldsymbol{a}^*)\right.
\end{aligned}$$

$$+\epsilon^2(a^j_{,j}a^{*l}_{,l} - a^l_{,j}a^{*j}_{,l})\Big] + \alpha^2\left(b + a^m\frac{\partial p_0}{\partial x^m}\right)\left(\epsilon a^{*j}_{,j} - \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a}^*\right)$$
$$+\alpha^2\left(b^* + a^{*m}\frac{\partial p_0}{\partial x^m}\right)\left(\epsilon a^j_{,j} + \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a}\right)\mathrm{O}(\alpha^4)\Big\}. \tag{6.1}$$

As in Section 5, the subscript with a comma denotes slow spatial derivatives, e.g., $a^i_{,j} \equiv \partial a^i/\partial\epsilon x^j$ and $\mathrm{d}/\mathrm{d}\epsilon t \equiv \partial/\partial\epsilon t + \bar{u}^j_{\mathrm{L}}\partial/\partial\epsilon x^j$.

The equation arising from stationarity of $\bar{\mathcal{L}}$ given by (6.1) under variations in p_0 at fixed $\boldsymbol{x}$ and t is

$$D = 1 - \alpha^2\epsilon^2(a^j a^{*l})_{,jl} + \mathrm{O}\alpha^4\epsilon) \tag{6.2}$$

which implies

$$\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot\bar{\boldsymbol{u}}_{\mathrm{L}} = \alpha^2\epsilon^2\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\left[(a^j a^{*l})_{,jl}\right] + \mathrm{O}(\alpha^4\epsilon). \tag{6.3}$$

Likewise, stationarity under variations in ϵl^A at fixed $\boldsymbol{x}$ and t implies the WMFI motion equation at this order, cf. Eq. (3.15),

$$\frac{\partial(\boldsymbol{m}/D)}{\partial\epsilon t} - \bar{\boldsymbol{u}}_{\mathrm{L}}\times\left(\frac{\partial}{\partial\epsilon\boldsymbol{x}}\times\frac{\boldsymbol{m}}{D}\right) + \frac{1}{\epsilon}\rho g\hat{z}$$
$$+\frac{\partial}{\partial\epsilon\boldsymbol{x}}\left[\frac{1}{2}|\bar{\boldsymbol{u}}_{\mathrm{L}}|^2 + \bar{p}_{\mathrm{L}} - \alpha^2\left(\frac{1}{D}\boldsymbol{p}\cdot\bar{\boldsymbol{u}}_{\mathrm{L}} + \mathcal{W}\right)\right] = \mathrm{O}(\alpha^4\epsilon). \tag{6.4}$$

The quantity $\alpha^2\mathcal{W}$ is $1/D$ times the (unconstrained) Lagrangian density due to the waves

$$\mathcal{W} = |\boldsymbol{a}|^2\tilde{\omega}^2 - \epsilon\mathrm{i}\tilde{\omega}\left(\boldsymbol{a}\cdot\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t} - \boldsymbol{a}^*\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right) + \epsilon^2\left|\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right|^2$$
$$+ 2\mathrm{i}\tilde{\omega}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*) + \epsilon\boldsymbol{\Omega}\cdot\left(\boldsymbol{a}\times\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t} + \boldsymbol{a}^*\times\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right) \tag{6.5}$$

and $\bar{p}_{\mathrm{L}}$ is the Lagrangian mean of the total pressure

$$\bar{p}_{\mathrm{L}} = \overline{p(\boldsymbol{x}+\alpha\boldsymbol{\xi}, t)} = p_0 + \alpha^2\frac{1}{2}\frac{\partial^2 p_0}{\partial x^i\partial x^j}(a^i a^{*j} + a^{*i}a^j) + \alpha^2\epsilon\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot(\boldsymbol{a}^*b + \boldsymbol{a}b^*) + \mathrm{O}(\alpha^4\epsilon). \tag{6.6}$$

The Eulerian momentum density $\boldsymbol{m}$ now takes the form, cf. Eq. (3.16),

$$\boldsymbol{m} = \frac{\delta\bar{\mathcal{L}}}{\delta\bar{\boldsymbol{u}}_{\mathrm{L}}} = D[\bar{\boldsymbol{u}}_{\mathrm{L}} + (\boldsymbol{\Omega}\times\boldsymbol{x})] - \alpha^2\boldsymbol{p}, \tag{6.7}$$

where $\boldsymbol{p}$ is the pseudomomentum density, given by, cf. [7],

$$\frac{p_j}{D} = -\overline{\frac{\partial\boldsymbol{\xi}}{\partial x^j}\cdot\left(\frac{\mathrm{d}\boldsymbol{\xi}}{\mathrm{d}t} + \boldsymbol{\Omega}\times\boldsymbol{\xi}\right)}$$
$$= k_j\frac{N}{D} - \epsilon\mathrm{i}\tilde{\omega}\left(\boldsymbol{a}^*\cdot\frac{\partial\boldsymbol{a}}{\partial\epsilon x^j} - \boldsymbol{a}\cdot\frac{\partial\boldsymbol{a}^*}{\partial\epsilon x^j}\right) - \epsilon\boldsymbol{\Omega}\cdot\left(\boldsymbol{a}\times\frac{\partial\boldsymbol{a}^*}{\partial\epsilon x^j} + \boldsymbol{a}^*\times\frac{\partial\boldsymbol{a}}{\partial\epsilon x^j}\right)$$
$$-\epsilon^2\left(\frac{\partial\boldsymbol{a}}{\partial\epsilon x^j}\cdot\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t} + \frac{\partial\boldsymbol{a}^*}{\partial\epsilon x^j}\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right) \tag{6.8}$$

and N is the wave action density, cf. Eq. (3.17),

$$N = \frac{1}{\alpha^2}\frac{\delta\overline{\mathcal{L}}}{\delta\omega} = D\left[2|\boldsymbol{a}|^2\tilde{\omega} - \epsilon\mathrm{i}\left(\boldsymbol{a}\cdot\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t} - \boldsymbol{a}^*\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right) + 2\mathrm{i}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*)\right]. \tag{6.9}$$

Within the decomposition (2.3), Eqs. (6.7)–(6.9) are exact.

Stationarity of $\overline{\mathcal{L}}$ given by (6.1) under variations in ϕ, b^*, and $\boldsymbol{a}^*$ at fixed $\boldsymbol{x}$ and t gives, respectively,

$$\frac{\partial N}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}}\cdot\left(\bar{\boldsymbol{u}}_{\mathrm{L}}N + \mathrm{i}D(\boldsymbol{a}^*b - \boldsymbol{a}b^*)\right) = \mathrm{O}(\alpha^2\epsilon), \tag{6.10}$$

$$\epsilon\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot\boldsymbol{a} + \mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a} = \mathrm{O}(\alpha^2\epsilon) \tag{6.11}$$

and, cf. Eq. (3.20),

$$\begin{aligned}&\tilde{\omega}^2\boldsymbol{a} + 2\epsilon\mathrm{i}\tilde{\omega}\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t} + \epsilon\mathrm{i}\boldsymbol{a}\frac{\mathrm{d}\tilde{\omega}}{\mathrm{d}\epsilon t} - \epsilon^2\frac{\mathrm{d}^2\boldsymbol{a}}{\mathrm{d}\epsilon t^2} + 2\left[\boldsymbol{\Omega}\times\left(\mathrm{i}\tilde{\omega}\boldsymbol{a} - \epsilon\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right)\right]\\ &-\epsilon\frac{\partial b}{\partial\epsilon\boldsymbol{x}} - \mathrm{i}b\boldsymbol{k} - \left(\boldsymbol{a}\cdot\frac{\partial}{\partial\boldsymbol{x}}\right)\frac{\partial p_0}{\partial\mathbf{x}} = \mathrm{O}(\alpha^2\epsilon).\end{aligned} \tag{6.12}$$

Thus, transversality of the wave amplitude vector $\boldsymbol{a}$ is broken at order $\mathrm{O}(\epsilon)$ and, at the same order, this vector becomes a prognostic variable, i.e., $\boldsymbol{a}$ acquires its own dynamical equation. Neglecting terms of order $\mathrm{O}(\epsilon)$ in Eqs. (6.2)–(6.12) recovers the equations of Section 3.2.

Looking toward the Hamiltonian formulation of these equations, it is convenient to introduce an auxiliary variable, in terms of which many of the expressions in this section simplify and become more compact. We introduce the momentum density $\alpha^2\boldsymbol{\pi}$, canonically conjugate to $\boldsymbol{a}^*$ and defined as

$$\boldsymbol{\pi} = \frac{1}{\alpha^2\epsilon}\frac{\delta\overline{\mathcal{L}}}{\delta(\partial\boldsymbol{a}^*/\partial\epsilon t)} = D\left(-\mathrm{i}\tilde{\omega}\boldsymbol{a} + \epsilon\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t} + \boldsymbol{\Omega}\times\boldsymbol{a}\right). \tag{6.13}$$

In terms of the canonical momentum $\boldsymbol{\pi}$, we find the following simpler forms, cf. Eqs. (6.5), (6.8), (6.9) and (6.12), respectively:

$$W = |\frac{\boldsymbol{\pi}}{D}|^2 - |\boldsymbol{\Omega}\times\boldsymbol{a}|^2, \tag{6.14}$$

$$p_j = Nk_j - 2\,\mathrm{Re}\left[\boldsymbol{\pi}\cdot\epsilon\frac{\partial\boldsymbol{a}^*}{\partial\epsilon x^j}\right], \tag{6.15}$$

$$N = -2\,\mathrm{Im}(\boldsymbol{\pi}\cdot\boldsymbol{a}^*), \tag{6.16}$$

$$\mathrm{i}\tilde{\omega}\boldsymbol{\pi} - \epsilon\frac{\mathrm{d}\boldsymbol{\pi}}{\mathrm{d}\epsilon t} + (\boldsymbol{\pi} - D\boldsymbol{\Omega}\times\boldsymbol{a})\times\boldsymbol{\Omega} - \epsilon D\frac{\partial b}{\partial\epsilon\boldsymbol{x}} - \mathrm{i}Db\boldsymbol{k} - D\left(\boldsymbol{a}\cdot\frac{\partial}{\partial\boldsymbol{x}}\right)\frac{\partial p_0}{\partial\boldsymbol{x}} = \mathrm{O}(\alpha^2\epsilon). \tag{6.17}$$

For a prescribed mean flow, the last equation, together with (6.11) and the wave transport equation, $\partial\boldsymbol{k}/\partial\epsilon t = \partial\omega/\partial\epsilon\boldsymbol{x}$, determines the linearized evolution of a wave train perturbation with velocity $\boldsymbol{v} = \mathrm{Re}(\boldsymbol{\pi}/D - \boldsymbol{\Omega}\times\boldsymbol{a})$ relative to the Lagrangian mean velocity $\bar{\boldsymbol{u}}_{\mathrm{L}}$. This is the connection between the present theory and the WKB wavepacket stability method. See, e.g., [33] for a clear exposition of the WKB stability method and references to its modern applications. See [13] for a recent application of the WKB stability method to the stability of the Kirchhoff ellipse in a rotating frame. However, unlike the WKB wavepacket equations discussed in [13,33], this restriction of the higher-order WMFI equations does not reduce to ordinary differential equations along fluid characteristics. Since the waves are not "frozen" into the fluid, the restricted WMFI equations remain partial differential equations along fluid characteristics.

Remark. Another self-consistent WMFI theory may be derived at an intermediate order, in which the term of order $O(\alpha^2\epsilon^2)$ is neglected in the averaged action in Eq. (6.1). Neglecting this term replaces $\boldsymbol{\pi}$ by $\boldsymbol{\pi}' = D(-i\tilde{\omega}\boldsymbol{a} + \boldsymbol{\Omega}\times\boldsymbol{a})$ and reduces the dynamical equations for $\boldsymbol{a}$ to first order, in accordance with the usual theory of wave propagation with slowly varying amplitudes. The equations of motion for this intermediate theory are summarized in Appendix A.

6.2. *Noether symmetries at higher order*

Invariance of $\overline{\mathcal{L}}$ in Eq. (6.1) under time translations leads to an energy conservation relation analogous to (4.17)

$$\frac{\partial W}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot \boldsymbol{S} = O(\alpha^4\epsilon), \tag{6.18}$$

where W is the total energy density, now given by

$$W/D = \tfrac{1}{2}|\bar{\boldsymbol{u}}_L|^2 + \alpha^2|\boldsymbol{\pi}/D - \boldsymbol{\Omega}\times\boldsymbol{a}|^2 + \rho g z + O(\alpha^4\epsilon), \tag{6.19}$$

where $\boldsymbol{\pi}$ is defined in Eq. (6.13) and $\boldsymbol{S}$ is the energy flux, which now takes the form

$$\begin{aligned} S^j = {}& \bar{u}^j_L(W + D\bar{p}_L) + \alpha^2 D\left[b\left(i\omega a^{*j} + \epsilon\frac{\partial a^{*j}}{\partial \epsilon t}\right) + b^*\left(-i\omega a^j + \epsilon\frac{\partial a^j}{\partial \epsilon t}\right)\right] \\ & + \alpha^2\epsilon\frac{\partial p_0}{\partial x^l}\frac{\partial}{\partial \epsilon t}(a^j a^{*l} + a^{*j}a^l) - \alpha^2\epsilon^2\frac{\partial}{\partial \epsilon x^l}\left[p_0\left(a^{*j}\frac{\partial a^l}{\partial \epsilon t} + a^j\frac{\partial a^{*l}}{\partial \epsilon t}\right)\right] + O(\alpha^4\epsilon). \end{aligned} \tag{6.20}$$

The momentum balance relation analogous to (4.21) is now

$$\frac{\partial G_i}{\partial \epsilon t} + \frac{\partial T^j_i}{\partial \epsilon x^j} + \frac{D}{\epsilon}\rho g\delta_{i3} + \frac{D}{\epsilon}(\boldsymbol{\Omega}\times\bar{\boldsymbol{u}}_L)_i = O(\alpha^4\epsilon), \tag{6.21}$$

where the total momentum density G_i is given by

$$G_i = m_i + \alpha^2 p_i = D[\bar{u}_{Li} + (\boldsymbol{\Omega}\times\boldsymbol{x})_i], \tag{6.22}$$

the wave pseudomomentum density p_i is given in Eq. (6.15), and the momentum stress tensor T^j_i at this order is given by

$$\begin{aligned} T^j_i = {}& \bar{u}^j_L G_i + D\bar{p}_L\delta^j_i - \alpha^2 D\left[b(-ik_i a^{*j} + \epsilon a^{*j}_{,i}) + b^*(ik_i a^j + \epsilon a^j_{,i})\right] \\ & - \alpha^2\epsilon\frac{\partial p_0}{\partial x^l}\frac{\partial}{\partial \epsilon x^i}(a^j a^{*l} + a^{*j}a^l) + \alpha^2\epsilon^2\frac{\partial}{\partial \epsilon x^l}\left[p_0(a^{*j}a^l_{,i} + a^j a^{*l}_{,i})\right] + O(\alpha^4\epsilon) \end{aligned} \tag{6.23}$$

with additional order $O(\epsilon\alpha^2)$ terms, relative to T^j_i in Eq. (4.23).

As for $\overline{\mathcal{L}}$ of Eq. (3.11), the averaged action at higher order in Eq. (6.1), is invariant under the transformation generated by $\epsilon^{ABC}\mu_C(\partial\rho/\partial l^B)(\partial/\partial l^A)$. Via an intermediate step analogous to Eq. (4.24) we obtain the local conservation relation

$$\frac{\partial Dq}{\partial \epsilon t} + \frac{\partial}{\partial \epsilon \boldsymbol{x}} \cdot Dq\bar{\boldsymbol{u}}_L = O(\alpha^4\epsilon), \tag{6.24}$$

where the potential vorticity q is defined by Eq. (4.26) but with $\boldsymbol{m}$ now explicitly taking the form (6.7) rather than (3.16). That is,

$$q = \frac{1}{D}\frac{\partial\rho}{\partial \boldsymbol{x}} \cdot \left[\frac{\partial}{\partial \epsilon \boldsymbol{x}} \times \left(\bar{\boldsymbol{u}}_L - \alpha^2\boldsymbol{p}/D + (\boldsymbol{\Omega}\times\boldsymbol{x})\right)\right] \tag{6.25}$$

with pseudomomentum density $\boldsymbol{p}$ given in Eq. (6.8). Use of the continuity equation (3.7) for D (which is still exact) then gives the local convection relation for q,

$$dq/\,d\epsilon t = O(\alpha^4\epsilon). \tag{6.26}$$

Remark on the Hamiltonian formulation of WMFI at higher order and the role of the pseudomomentum density. The Hamiltonian formulation of the higher-order WMFI theory discussed in this section has the *same* Lie–Poisson Hamiltonian structure as for the leading order WMFI dynamics discussed in Section 5, but with a new Eulerian momentum density $\boldsymbol{m}$ defined by Eq. (6.7) instead of Eq. (3.16), and with $\boldsymbol{a}^*$ and $\alpha^2\pi$ as new canonically conjugate wave variables, in addition to ϕ/ϵ and $-\alpha^2 N$. Thus, the Hamiltonian formulation of the higher-order WMFI theory consists of a Poisson bracket which is the sum of the usual Lie–Poisson bracket for the fluid variables, $(\boldsymbol{m}, D, \rho)$, with redefined $\boldsymbol{m}$, plus a canonical Poisson bracket for the wave variables, $(-\alpha^2 N, \phi/\epsilon)$ and $(\alpha^2\boldsymbol{\pi}, \boldsymbol{a}^*)$. In contrast to the phase variable ϕ/ϵ, however, the amplitude variable $\boldsymbol{a}^*$ is not ignorable, and thus $\alpha^2\pi$ does not satisfy a conservation law, cf. Eq. (6.17). The wave and mean-flow variables are coupled through their (constrained) Hamiltonian and possess the total momentum density, $D(\bar{\boldsymbol{u}}_L + (\boldsymbol{\Omega}\times\boldsymbol{x})) = \boldsymbol{m} + \alpha^2\boldsymbol{p}$. See [22] for full details.

6.3. Prescribed fluctuations

6.3.1. GLM theory

The present results can be restricted to recover the GLM theory of Andrews and McIntyre [7] by taking the rapid fluctuations as prescribed. This restriction of the WMF theory provides a variational formulation of the GLM theory. The GLM theory follows upon replacing the decomposition of the pressure (2.11) by

$$p(\boldsymbol{x}+\alpha\boldsymbol{\xi}, t) = \bar{p}_L(\boldsymbol{x}, t) + \sum_{j=1}^{3} \alpha^j \left(h_j e^{ij\phi/\epsilon} + h_j^* e^{-ij\phi/\epsilon}\right), \tag{6.27}$$

and then assuming that the rapidly fluctuating displacement $\boldsymbol{\xi}$ is a *prescribed* function of $\boldsymbol{x}$ and t, which satisfies the transversality condition (6.29) below. Since no variations of the averaged action are to be taken with respect to $\boldsymbol{a}$ or ϕ, it is not necessary to use the explicit form of $\boldsymbol{\xi}$ given by (2.5). The averaged action is

$$\begin{aligned}\bar{\mathcal{L}} = \int dt \int d^3x \Bigg\{ & D\Bigg[\frac{1}{2}|\bar{\boldsymbol{u}}_L|^2 + \frac{\alpha^2}{2}\overline{\left|\frac{\partial\boldsymbol{\xi}}{\partial t}\right|^2} + \alpha^2 \bar{u}_L^i \overline{\frac{\partial\boldsymbol{\xi}}{\partial x^i}\cdot\frac{\partial\boldsymbol{\xi}}{\partial t}} + \frac{\alpha^2}{2}\bar{u}_L^i\bar{u}_L^j \overline{\frac{\partial\boldsymbol{\xi}}{\partial x^i}\cdot\frac{\partial\boldsymbol{\xi}}{\partial x^j}} \\ & -\rho\left(l^A(\boldsymbol{x}, t)\right) gz + \bar{\boldsymbol{u}}_L\cdot(\boldsymbol{\Omega}\times\boldsymbol{x}) + \alpha^2 \overline{\frac{\partial\boldsymbol{\xi}}{\partial t}\cdot(\boldsymbol{\Omega}\times\boldsymbol{\xi})} + \alpha^2\bar{u}_L^i \overline{\frac{\partial\boldsymbol{\xi}}{\partial x^i}\cdot(\boldsymbol{\Omega}\times\boldsymbol{\xi})}\Bigg] \\ & + \bar{p}_L\left[1 - D + \frac{\alpha^2}{2}\frac{\partial}{\partial x^i}\overline{\left(\xi^i\frac{\partial\xi^j}{\partial x^j} - \xi^j\frac{\partial\xi^i}{\partial x^j}\right)}\right] \\ & + \alpha^2\left[h_1 \overline{e^{i\phi/\epsilon}\frac{\partial\xi^j}{\partial x^j}} + h_1^* \overline{e^{-i\phi/\epsilon}\frac{\partial\xi^j}{\partial x^j}}\right] + O(\alpha^4)\Bigg\}.\end{aligned} \tag{6.28}$$

The neglected terms are independent of l^A and its derivatives and of $\bar{p}_L$.

The variation of $\bar{\mathcal{L}}$ of Eq. (6.28) with respect to h_1 at fixed $\boldsymbol{x}$ and t yields

$$\overline{e^{i\phi/\epsilon}\frac{\partial\xi^j}{\partial x^j}} = O(\alpha^2\epsilon), \tag{6.29}$$

which with the form of $\boldsymbol{\xi}$ given in (2.5) can be written as, cf. Eq. (6.11),

$$\epsilon\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot\boldsymbol{a}+\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a}=\mathrm{O}(\alpha^2\epsilon). \tag{6.30}$$

Likewise, the variation of $\overline{\mathcal{L}}$ with respect to $\overline{p}_{\mathrm{L}}$ at fixed $\boldsymbol{x}$ and t gives

$$D=1+\frac{\alpha^2}{2}\frac{\partial}{\partial x^i}\overline{\left(\xi^i\frac{\partial\xi^j}{\partial x^j}-\xi^j\frac{\partial\xi^i}{\partial x^j}\right)}, \tag{6.31}$$

and the variation with respect to l^A results in the GLM motion equation,

$$\begin{aligned}&\frac{\partial(\boldsymbol{m}/D)}{\partial\epsilon t}-\overline{\boldsymbol{u}}_{\mathrm{L}}\times\left(\frac{\partial}{\partial\epsilon\boldsymbol{x}}\times\frac{\boldsymbol{m}}{D}\right)+\frac{1}{\epsilon}\rho g\hat{\boldsymbol{z}}\\&+\frac{\partial}{\partial\epsilon\boldsymbol{x}}\left[\overline{p}_{\mathrm{L}}+|\overline{\boldsymbol{u}}_{\mathrm{L}}|^2-\frac{1}{2}\overline{\boldsymbol{u}^{\xi}\cdot\boldsymbol{u}^{\xi}}-\overline{\boldsymbol{u}^{\xi}\cdot(\boldsymbol{\Omega}\times\boldsymbol{\xi})}-\alpha^2\boldsymbol{p}\cdot\overline{\boldsymbol{u}}_{\mathrm{L}}/D\right]=0,\end{aligned} \tag{6.32}$$

where $\boldsymbol{m}$ and $\boldsymbol{p}$ are defined in Eqs. (6.7) and (6.8), respectively. Here we have adopted the notation of Andrews and McIntyre [7] $\boldsymbol{u}^{\xi}=\boldsymbol{U}(\boldsymbol{x}+\alpha\boldsymbol{\xi},t)$ to write

$$\tfrac{1}{2}|\overline{\boldsymbol{u}}_{\mathrm{L}}|^2-\alpha^2\mathcal{W}=|\overline{\boldsymbol{u}}_{\mathrm{L}}|^2-\tfrac{1}{2}\overline{\boldsymbol{u}^{\xi}\cdot\boldsymbol{u}^{\xi}}-\overline{\boldsymbol{u}^{\xi}\cdot(\boldsymbol{\Omega}\times\boldsymbol{\xi})}. \tag{6.33}$$

Formally, no truncation in α or ϵ has been effected in Eqs. (6.31) and (6.32). Using the definition of $\boldsymbol{\xi}$ in Eq. (2.5) and the transversality condition (6.30), the constraint (6.31) reduces to Eq. (6.2). Eqs. (6.29), (6.31) and (6.32) are the equations of the GLM theory for incompressible flow discussed in [7]. Not unexpectedly, the Hamiltonian structure for the higher-order WMFI equations of the previous section formally reproduces these GLM equations when the wave displacement $\boldsymbol{\xi}(\boldsymbol{x},t)$ is a prescribed quantity, see [22].

Cautionary note. In view of our earlier discussion in Section 2.4, in the case of prescribed fluctuations the averaging theorem does not lead to Eqs. (2.17), (2.18) and (2.27) (or, at higher order, to Eqs. (6.2), (6.4) and (6.11) or (6.29), (6.31) and (6.32)), unless the prescribed fluctuations happen to satisfy the projection conditions in Eqs. (2.28) and (2.29) (or, at higher order, (6.10) and (6.12)) relating the amplitude, phase, and phase derivatives of the fluctuations. If these conditions are not met by the prescribed fluctuations, the decomposition into mean and prescribed fluctuating quantities will not be preserved in time.

Acknowledgements

We are grateful for valuable discussions with A. Brizzard, L. Bunimovich, P. Constantin, R. Grimshaw, A.N. Kaufman, A. Lifschitz (Lipton), L.G. Margolin, J.C. McWilliams and V.E. Zakharov. This work is supported by the US Department of Energy.

Appendix A. WMFI equations at intermediate order and conservation of wave angular momentum

Neglecting terms of order $\mathrm{O}(\alpha^2\epsilon^2)$ in Section 6.1 and suppressing the asymptotic order estimates yields the following self-consistent WMFI equations (with $D=1$):

$$\begin{aligned}&\left(\frac{\partial}{\partial\epsilon t}+\overline{\boldsymbol{u}}_{\mathrm{L}}\cdot\frac{\partial}{\partial\epsilon\boldsymbol{x}}\right)\overline{\boldsymbol{u}}_{\mathrm{L}}-\frac{1}{\epsilon}\left(\overline{\boldsymbol{u}}_{\mathrm{L}}\times2\boldsymbol{\Omega}-\rho g\hat{\boldsymbol{z}}-\frac{\partial p_0}{\partial\boldsymbol{x}}\right)-\alpha^2\left[\left(\frac{\partial}{\partial\epsilon t}+\overline{\boldsymbol{u}}_{\mathrm{L}}\cdot\frac{\partial}{\partial\epsilon\boldsymbol{x}}\right)\boldsymbol{p}+p_j\frac{\partial}{\partial\epsilon\boldsymbol{x}}\overline{u}^j_{\mathrm{L}}\right]\\&+\frac{\alpha^2\epsilon}{2}\frac{\partial}{\partial\epsilon\boldsymbol{x}}\left[\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot(\boldsymbol{a}\boldsymbol{b}^*+\boldsymbol{a}^*\boldsymbol{b})\right]=0\end{aligned} \tag{A.1}$$

with

$$\frac{\mathrm{d}\rho}{\mathrm{d}t}=0,\quad \frac{\partial N}{\partial\epsilon t}+\frac{\partial}{\partial\epsilon \boldsymbol{x}}\cdot\big(\overline{\boldsymbol{u}}_{\mathrm{L}}N+\mathrm{i}(\boldsymbol{a}^*b-\boldsymbol{a}b^*)\big)=0,\quad \frac{\partial \boldsymbol{k}}{\partial\epsilon t}+\frac{\partial\omega}{\partial\epsilon\boldsymbol{x}}=0. \tag{A.2}$$

The quantities N and $\boldsymbol{p}$ are defined by

$$N=\frac{1}{\alpha^2}\frac{\delta\overline{\mathcal{L}}}{\delta\omega}=2|\boldsymbol{a}|^2\tilde{\omega}-\epsilon\mathrm{i}\left(\boldsymbol{a}\cdot\frac{\mathrm{d}\boldsymbol{a}^*}{\mathrm{d}\epsilon t}-\boldsymbol{a}^*\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right)+2i\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*) \tag{A.3}$$

and

$$p_j=k_jN+2\,\mathrm{Re}\left[\epsilon\frac{\partial\boldsymbol{a}^*}{\partial\epsilon x^j}\cdot(\mathrm{i}\tilde{\omega}\boldsymbol{a}-\boldsymbol{\Omega}\times\boldsymbol{a})\right]. \tag{A.4}$$

We recall from Eq. (4.14) that the Stokes mean drift velocity is given by the similar, but different, expression

$$\overline{\boldsymbol{u}}_{\mathrm{S}}=\alpha^2\epsilon 2\,\mathrm{Re}\frac{\partial}{\partial\epsilon x^i}(\mathrm{i}\tilde{\omega}a^i\boldsymbol{a}^*)+\mathrm{O}(\alpha^2\epsilon^2,\alpha^4). \tag{A.5}$$

So the pseudomomentum density $\alpha^2\boldsymbol{p}$ and the Stokes mean drift velocity $\overline{\boldsymbol{u}}_{\mathrm{S}}$ are different quantities of different magnitudes which cannot be mistaken for each other.

The diagnostic variables, p_0, b, and $\tilde{\omega}=\omega-\overline{\boldsymbol{u}}_{\mathrm{L}}\cdot\boldsymbol{k}$ and the prognostic variable $\boldsymbol{a}$ are determined from the constraints (again with $D=1$)

$$\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot\overline{\boldsymbol{u}}_{\mathrm{L}}=0,\qquad \epsilon\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot\boldsymbol{a}+\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a}=0 \tag{A.6}$$

and the dynamical equation for the vector wave amplitude, $\boldsymbol{a}$,

$$\tilde{\omega}^2\boldsymbol{a}+2\mathrm{i}\tilde{\omega}(\boldsymbol{\Omega}\times\boldsymbol{a})-\mathrm{i}\boldsymbol{k}b-\left(\boldsymbol{a}\cdot\frac{\partial}{\partial\boldsymbol{x}}\right)\frac{\partial p_0}{\partial\boldsymbol{x}}+\epsilon\left(2\mathrm{i}\tilde{\omega}\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}+i\boldsymbol{a}\frac{\mathrm{d}\tilde{\omega}}{\mathrm{d}\epsilon t}-2\boldsymbol{\Omega}\times\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}-\frac{\partial b}{\partial\epsilon\boldsymbol{x}}\right)=0. \tag{A.7}$$

These equations result from HP with action $\overline{\mathcal{L}}$ in (6.1) truncated at order $\mathrm{O}(\epsilon^2\alpha^2)$ to remove *only* the term $\epsilon^2\alpha^2|\,\mathrm{d}\boldsymbol{a}/\,\mathrm{d}\epsilon t|^2$ from its integrand. The truncated HP at this order is invariant under translations in phase ϕ because of averaging, as usual. However, it is also invariant under the restricted phase shift (3.22). By Noether's theorem, this restricted phase invariance implies that the quantity

$$N'=D\big(2|\boldsymbol{a}|^2\tilde{\omega}+2\mathrm{i}\boldsymbol{\Omega}\cdot(\boldsymbol{a}\times\boldsymbol{a}^*)\big), \tag{A.8}$$

is a conserved density, satisfying, cf. Eq. (A.2),

$$\frac{\partial N'}{\partial\epsilon t}+\frac{\partial}{\partial\epsilon\boldsymbol{x}}\cdot\big(N'\overline{\boldsymbol{u}}_{\mathrm{L}}+\mathrm{i}(\boldsymbol{a}^*b-\boldsymbol{a}b^*)\big)=0. \tag{A.9}$$

The *same* conservation law can be obtained by taking the imaginary part of $\boldsymbol{a}^*\cdot$ (A.7). Therefore the difference of the conservation laws for N and N' gives a new WMFI conservation law at this intermediate truncation order. Indeed, this difference gives

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon t}\mathrm{Im}\left(\boldsymbol{a}^*\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right)=0 \tag{A.10}$$

which is conservation of $(N-N')/D$ on fluid parcels. Writing $a_k=r_k\exp(\mathrm{i}\theta_k)$, $k=1,2,3$ (no sum) expresses the quantity

$$\mathrm{Im}\left(\boldsymbol{a}^*\cdot\frac{\mathrm{d}\boldsymbol{a}}{\mathrm{d}\epsilon t}\right)=\sum_{k=1}^{3}r_k^2\frac{\mathrm{d}\theta_k}{\mathrm{d}\epsilon t}, \tag{A.11}$$

as a "wave angular momentum" which by Eq. (A.10) is embedded in the mean flow. In the lowest order WMFI theory discussed in Section 3.2, invariance of $\overline{\mathcal{L}}$ in (3.11) under the restricted phase shift (3.22) has trivial consequences – its conserved density is a pure gauge. At this intermediate order, invariance of $\overline{\mathcal{L}}$ in (6.1) truncated at order $O(\alpha^2\epsilon^2)$ under the restricted phase shift produces the new convection law (A.10) for wave angular momentum. In the highest order WMFI theory of Section 6, invariance of $\overline{\mathcal{L}}$ in (6.1) under the restricted phase shift produces the *same* result as translation invariance in ϕ does, namely the wave action conservation law for N in (6.9) or (A.3) and the wave angular momentum convection law (A.10) is lost at that order.

One approach to the solution of this system at the intermediate truncation order is to expand all variables in powers of ϵ and solve using matched asymptotics. For convenience we write $\boldsymbol{a}$ in the form

$$\boldsymbol{a}(\epsilon \boldsymbol{x}, \epsilon t) = a(\epsilon \boldsymbol{x}, \epsilon t)\hat{\boldsymbol{e}}(\epsilon \boldsymbol{x}, \epsilon t), \tag{A.12}$$

where the displacement magnitude $a(\epsilon \boldsymbol{x}, \epsilon t)$ is real. Then, at order O(1), Eq. (A.7) determines the (unit) polarization vector $\hat{\boldsymbol{e}}$, but leaves the magnitude a and the overall phase of $\hat{\boldsymbol{e}}$ undetermined. The magnitude a is determined by the initial value of the conserved density N' and its subsequent evolution, cf. Eq. (A.9). The overall phase of $\hat{\boldsymbol{e}}$, on the other hand, is determined from the initial value of $\mathrm{Im}(\boldsymbol{a}^* \cdot \mathrm{d}\boldsymbol{a}/\mathrm{d}\epsilon t)$, which is preserved on fluid parcels at this intermediate truncation order.

References

[1] H.D.I. Abarbanel and D.D. Holm, Nonlinear stability of inviscid flows in three dimensions: incompressible fluids and barotropic fluids, Phys. Fluids 30 (1987) 3369–3382.

[2] H.D.I. Abarbanel, D.D. Holm, J.E. Marsden and T. Ratiu, Nonlinear stability analysis of stratified ideal fluid equilibria, Philos. Trans. Roy. Soc. London Ser. A 318 (1987) 349–409.

[3] D.G. Andrews, A conservation law for small-amplitude quasi-geostrophic disturbances on a zonally asymmetric basic flow, J. Atmospheric Sci. 40 (1983) 85–90.

[4] D.G. Andrews, A finite-amplitude Eliassen–Palm theorem in isentropic coordinates, J. Atmospheric Sci. 40 (1983) 1877–1883.

[5] D.G. Andrews, On the interpretation of the Eliassen–Palm flux divergence, J.R. Met. Soc. 113 (1987) 323–338.

[6] D.G. Andrews, J.R. Holton and C.B. Leovy, Middle Atmosphere Dynamics (Academic Press, New York, 1987).

[7] D.G. Andrews and M.E. McIntyre, An exact theory of nonlinear waves on a Lagrangian-mean flow, J. Fluid Mech. 89 (1987) 609–646.

[8] D.G. Andrews and M.E. McIntyre, On wave-action and its relatives, J. Fluid Mech. 89 (1978) 647–664; addendum ibid 95, 796.

[9] V.I. Arnold, Sur la géometrie differentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits, Ann. Inst. Fourier (Grenoble) 16 (1966) 319–361.

[10] V.I. Arnold, On an a priori estimate in the theory of hydrodynamical stability, Am. Math. Soc. Transl. Ser. (2) 79 (1966) 267–269.

[11] V.I. Arnold, The Hamiltonian nature of Euler's equations in dynamics of a rigid body and a perfect fluid, Uspekhi Mat. Nauk. 24 (1969) 225–226 (in Russian).

[12] G.K. Batchelor, An Introduction to Fluid Mechanics (Cambridge University Press, London, 1985).

[13] B.J. Bayly, D.D. Holm and A. Lifschitz, Three-dimensional stability of elliptical vortex columns in external strain flows, Trans. Roy. Soc. London A 354 (1996) 895–926.

[14] F.P. Bretherton, On the mean motion induced by internal gravity waves, J. Fluid Mech. 89 (1969) 609–646.

[15] F.P. Bretherton, The general linearized theory of wave propagation, in: Mathematical Problems in the Geophysical Sciences, Lectures in Applied Mathematics, Vol. 13 (A.M.S. Providence, RI, 1971) pp. 61–102.

[16] F.P. Bretherton and C.J.R. Garrett, Wavetrains in inhomogeneous moving media, Proc. Roy. Soc. A 302 (1969) 529–554.

[17] J.G. Charney and P.G. Drazin, Propagation of planetary-scale disturbances from the lower into the upper atmosphere, J. Geophys. Res. 66 (1961) 83–109.

[18] R.L. Dewar, Interaction between hydromagnetic waves and a time-dependent inhomogeneous medium, Phys. Fluids 13 (1970) 2710–2720.

[19] D.G.B. Edelen, Aspects of variational arguments in the theory of elasticity: fact and folklore, Int. J. Eng. Solids Structures 17 (1981) 729–740.

[20] D.C. Fritts, S. Sun and D.-Y. Wang, Wave-wave interactions in a compressible atmosphere 1. A general formulation including rotation and wind shear, J. Geophys. Res. 97 (D9) (1992) 9975–9988.

[21] C. Garrett and W. Munk, Internal waves in the ocean, Ann. Rev. Fluid Mech. 11 (1979) 339–369.

[22] I. Gjaja and D.D. Holm, Self-consistent Hamiltonian formulation of wave mean-flow interaction for a rotating stratified incompressible fluid, in preparation.
[23] R. Grimshaw, Nonlinear internal gravity waves in a rotating fluid, J. Fluid Mech. 71 (1975) 479–512.
[24] R. Grimshaw, Nonlinear internal gravity waves and their interaction with the mean wind, J. Atmospheric Sci. 32 (1975) 1779–1793.
[25] R. Grimshaw, Mean flows induced by internal gravity wave packets propagating in a shear flow, Philos. Trans. Roy. Soc. A 292 (1979) 391–417.
[26] R. Grimshaw, Wave action and wave-mean flow interaction with application to stratified shear flows, Ann. Rev. Fluid Mech. 16 (1984) 11–44.
[27] D.D. Holm, The ideal Craik–Leibovich equations, Physica D 98 (1996) 415–441, these proceedings.
[28] D.D. Holm and B.A. Kupershmidt, Poisson structures of superfluids, Phys. Lett. A 91 (1982) 425–430.
[29] D.D. Holm and B.A. Kupershmidt, Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity, Physica D 6 (1983) 347–363.
[30] D.D. Holm, B. Kupershmidt and C.D. Levermore, Canonical maps between Poisson brackets in Eulerian and Lagrangian descriptions of continuum mechanics, Phys. Lett. A 98 (1983) 389–395.
[31] D.D. Holm, J.E. Marsden and T. Ratiu, Hamiltonian Structure and Lyapunov Stability for Ideal Continuum Dynamics, ISBN 2-7606-0771-2 (University of Montreal Press, Montreal, 1987).
[32] D.D. Holm, J.E. Marsden, T. Ratiu and A. Weinstein, Nonlinear stability of fluid and plasma equilibria, Phys. Rep. 123 (1985) 1–116.
[33] A. Lifschitz, On the instability of certain motions of an ideal incompressible fluid, Adv. Appl. Math. 15 (1994) 404–436.
[34] P. Lochak and C. Meunier, Multiphase Averaging for Classical Systems (Springer, New York, 1988).
[35] J.E. Marsden, T. Ratiu and A. Weinstein, Semi-direct products and reduction in mechanics, Trans. Amer. Math. Soc. 281 (1984) 147–177.
[36] M.E. McIntyre and T.G. Shepherd, An exact local conservation theorem for finite-amplitude disturbances to nonparallel shear flows, with remarks on Arnol'd stability theorems, J. Fluid Mech. 181 (1987) 527–565.
[37] M.J. McPhaden and P. Ripa, Wave-mean flow interactions in the equatorial ocean, Ann. Rev. Fluid Mech. 22 (1990) 167–205.
[38] P. Müller, G. Holloway, F. Henyey and N. Pomphrey, Nonlinear interactions among internal gravity waves, Rev. Geophys. 24 (1986) 493–536.
[39] P.J. Olver, Applications of Lie groups to Differential Equations, 2nd Ed. (Springer, New York, 1993).
[40] J. Pedlosky, Geophysical Fluid Dynamics, 2nd Ed. (Springer, New York, 1987).
[41] S.J. Putterman, P.H. Roberts, Non-linear hydrodynamics and a one-fluid theory of superfluid He^4, Phys. Lett. A 89 (1982) 444–447.
[42] E.J. Routh, A. Treatise on the Dynamics of a System of Rigid Bodies (Dover, New York, 1955).
[43] R. Salmon, Hamiltonian fluid mechanics, Ann. Rev. Fluid Mech. 20 (1988) 225–256.
[44] J.A. Sanders and F. Verhulst, Averaging Methods in Nonlinear Dynamical Systems (Springer, New York, 1985).
[45] T.G. Shepherd, Symmetries, conservation laws and Hamiltonian structure in geophysical fluid dynamics, Adv. Geophys. 32 (1990) 287–338.
[46] G.B. Whitham, A general approach to linear and nonlinear dispersive waves using a Lagrangian, J. Fluid Mech. 22 (1965) 273–283.
[47] G.B. Whitham, Two-timing, variational principles and waves, J. Fluid Mech. 44 (1970) 373–395.